The variation of the sum of edge lengths in linear arrangements of trees

Ramon Ferrer-i-Cancho[1], **Carlos Gómez-Rodríguez**[2] **and Juan Luis Esteban**[3]

[1] Complexity & Quantitative Linguistics Lab, LARCA Research Group
Departament de Ciències de la Computació,
Universitat Politècnica de Catalunya,
Campus Nord, Edifici Omega, Jordi Girona Salgado 1-3.
08034 Barcelona, Catalonia (Spain)
[2] Universidade da Coruña, CITIC
FASTPARSE Lab, LyS Research Group
Departamento de Ciencias de la Computación y Tecnologías de la Información,
Facultade de Informática,
Campus de A Coruña, 15071 A Coruña, Spain
[3] Departament de Ciències de la Computació,
Universitat Politècnica de Catalunya,
Campus Nord, Edifici Omega, Jordi Girona Salgado 1-3.
08034 Barcelona, Catalonia (Spain)

E-mail: `rferrericancho@cs.upc.edu, cgomezr@udc.es, esteban@cs.upc.edu`

Abstract. A fundamental problem in network science is the normalization of the topological or physical distance between vertices, that requires understanding the range of variation of the unnormalized distances. Here we investigate the limits of the variation of the physical distance in linear arrangements of the vertices of trees. In particular, we investigate various problems on the sum of edge lengths in trees of a fixed size: the minimum and the maximum value of the sum for specific trees, the minimum and the maximum in classes of trees (bistar trees and caterpillar trees) and finally the minimum and the maximum for any tree. We establish some foundations for research on optimality scores for spatial networks in one dimension.

PACS numbers: 89.75.Hc Networks and genealogical trees
89.75.Da Systems obeying scaling laws
89.75.Fb Structures and organization in complex systems

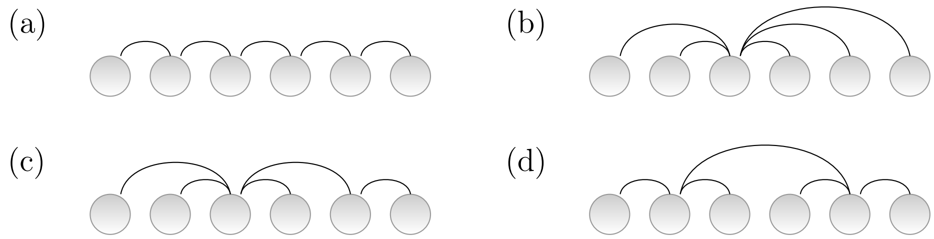


Figure 1. Different trees with $n = 6$ vertices and maximum degree k_1. (a) Linear tree ($k_1 = 2$). (b) Star tree ($k_1 = n - 1 = 5$). (c) Quasistar tree ($k_1 = n - 2 = 4$). (d) Balanced bistar tree ($k_1 = \lceil n/2 \rceil = 3$), that is formed by two star trees of 3 vertices joined by their respective hubs.

1. Introduction

A fundamental problem in network science is the normalization of the distance between vertices [1, 2, 3, 4, 5]. The problem is actually two-fold depending on whether focus is on topological distance (the distance between vertices in terms of number of edges) [1, 4, 5] or physical distance (the distance between vertices in some metric space, that may not be Euclidean) [2, 3].

Concerning topological distance, namely distance on a network, the simplest measure of topological distance is the average path length or characteristic path length, which can be defined on an undirected network g of n vertices as [6],

$$\langle l \rangle^g = \frac{1}{\binom{n}{2}} \sum_{i<j} l_{ij},$$

where l_{ij} is the minimum distance in edges between vertices i and j in g. In a connected network, $\langle l \rangle^g$ varies between its value in a complete graph, a graph with as many edges as possible, and a linear tree, a tree with maximum degree two (figure 1 (a)), i.e. [4]

$$\langle l \rangle^{complete} \leq \langle l \rangle \leq \langle l \rangle^{linear},$$

where

$$\langle l \rangle^{complete} = 1$$
$$\langle l \rangle^{linear} = \frac{n+1}{3}.$$

In trees, connected networks minimizing the number of edges, one has [4]

$$\langle l \rangle^{star} \leq \langle l \rangle^g \leq \langle l \rangle^{linear}, \tag{1}$$

where

$$\langle l \rangle^{star} = 2(n-1)/n$$

corresponds to a star tree, a tree with a hub of maximum degree, namely $n - 1$ (figure 1 (b)). In [4], the ratio

$$\lambda^g = \frac{\langle l \rangle^g}{\langle l \rangle^{linear}},$$

was used as a normalized measure of topological distance cost ($\lambda \leq 1$). Recently, two normalizations of $\langle l \rangle$ have been investigated [5]

$$\lambda^g_{\prime} = \frac{\langle l \rangle^g}{\langle l \rangle^g_{US}}$$

$$\lambda^g_{\prime\prime} = \frac{\langle l \rangle^g - \langle l \rangle^g_{US}}{\langle l \rangle^g_{UL} - \langle l \rangle^g_{US}},$$

where $\langle l \rangle^g_{US}$ and $\langle l \rangle^g_{UL}$ are the minimum (Ultra Short) and the maximum (Ultra Long) value of $\langle l \rangle$ of a network with same number of vertices and edges. In a tree, $\langle l \rangle^g_{US} = \langle l \rangle^{star}$ and $\langle l \rangle^g_{UL} = \langle l \rangle^{linear}$.

A limitation of $\langle l \rangle$ is that $l_{ij} = \infty$ for vertices in different connected components and then $\langle l \rangle$ is not finite in disconnected graphs, regardless of how closely connected vertices are within each component [6]. For this reason, an alternative is the so-called network *efficiency* [1], an average of $1/l_{ij}$ that can be defined as

$$E^g = \frac{1}{\binom{n}{2}} \sum_{i<j} \frac{1}{l_{ij}}.$$

$1/E^g$ is a harmonic mean and E is an average that is already normalized, since

$$E^{empty} \leq E^g \leq E^{complete}.$$

$E^{empty} = 0$ and $E^{complete} = 1$ correspond to an empty network (a network with no edges) and a complete graph, respectively. E^g is a normalized measure of optimality of a network with respect to topological distance.

In relation to physical distance, distance is often measured as D^g, the sum of the lengths of all links of g, where the length of a link is the physical distance between the vertices [3]. In a network of m edges, the average edge length [2]

$$\langle d \rangle^g = \frac{1}{m} D^g,$$

is the counterpart of $\langle l \rangle^g$ in physical space. For reference, D^g is compared against D^g_{MST}, the value of D of a minimum spanning tree of the original graph, namely a tree t defined on a subset of the edges of the original graphs such that D^t is minimum [3]. As $D^g_{MST} \leq D^g$, a normalized measure of physical distance cost is [3]

$$C^g = \frac{D^g}{D^g_{MST}},$$

namely a measure the degree of optimality of a network from the perspective of the topological distance.

Traditionally, D^g has been defined on a Euclidean two-dimensional space [3]. Here we focus on the problem of the range of variation of the physical distance D^g in one dimension, in particular in linear arrangements of the vertices. A prototypical example is the syntactic dependency network of a sentence, where vertices are words, edges indicate syntactic dependencies and the order of the words the sentence defines a linear arrangement (figure 2) [7]. There the length of an edge is usually defined as the absolute value of the difference between vertex positions: then consecutive words are at distance

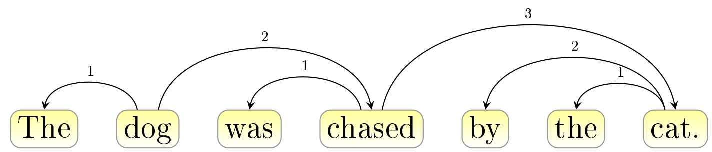


Figure 2. A syntactic dependency structure of a sentence adapted from `https://universaldependencies.org/introduction.html`. Here edge labels indicate the distance between the linked words (in words).

1, words separated by a word are at distance 2 and so on [2]. In the example (figure 2), $D^g = 10$. We are interested in the variation of D^g when the network structure remains constant, i.e. the limits of the variation of D over the $n!$ linear arrangements. Given a network g, the calculation D^g_{min}, the minimum value of D^g over all linear arrangements is known as the minimum linear arrangement problem [8], whereas the calculation of the maximum, i.e. D^g_{max}, is known as the maximum linear arrangement problem [9]. Both problems are computationally hard [8, 9]. In a tree t, the minimum linear arrangement problem simplifies and can be computed in polynomial time [10, 11, 12] but still formulae for D^t_{min} and D^t_{max} are only available for specific trees [13, 12, 14].

D^t_{min} and D^t_{max} and their limits of variation are relevant for research on the efficiency of language, where various optimality scores have been considered [2, 15, 16]. The first optimality score for D^t that was defined is [2, 15]

$$\Gamma^t = \frac{D^t}{D^t_{min}}.$$

Γ^t is the analog of the physical distance cost C^g for research on D^g where g is a fixed tree t and D^t varies depending only on the linear arrangement. Another score that has been considered is [16]

$$\Delta^t = D^t - D^t_{min}.$$

These limits are also relevant for a recently introduced z-scored value of D, i.e. [17]

$$D^t_z = \frac{D - D_{rla}}{(\mathbb{V}^t_{rla})^{1/2}}, \tag{2}$$

where D_{rla} and $\mathbb{V}^t_{rla}$ are, respectively, the expectation and the variance of D^t in a uniformly random linear arrangement (r.l.a.). D_{rla} depends only on n, as [18]

$$\mathbb{E}_{rla}[D^t] = \frac{1}{3}(n^2 - 1). \tag{3}$$

Given the potential to obtain simple formulae for trees and the interest of trees in language research [7, 19], here we are interested in three kinds of problems over trees of n vertices:

(i) D^t_{min} and D^t_{max} in specific kinds of trees (distinct unlabelled trees) that are selected for their theoretical interest. Linear trees and star trees are relevant to understand the variation of topological distance as we have seen above (equation 1) [4] and also to understand the limits of the variation of D^t_{min} [13, 20]. Trivially [12, 14, 20],

$$D^{linear}_{min} = n - 1.$$

Iordanskii found that [13],

$$D^{star}_{min} = \left\lfloor \frac{1}{4} n^2 \right\rfloor,$$

which was rediscovered later in equivalent forms [14, 20], e.g.

$$D^{star}_{min} = \frac{1}{4}(n^2 - n \bmod 2).$$

Bistar trees (*bistar*) consist of two stars joined by the hub and include star graphs as an extreme case when one of the stars has only one vertex (figures 1 (b-d)) [21, 22]. Here we are interested in two distinct representatives: quasistar trees (*quasi*), where one of the original stars has only two vertices (figure 1 (c)) and balanced bistar trees (*b-bistar*), where the two original stars have the same size or differ in one vertex (figure 1 (d)). Quasistar trees are important for the theory of edge crossings in linear arrangements [23, 24]. In this article, we will unveil that balanced bistar trees maximize D^t_{max} over trees of n vertices. Here we will obtain formulae for D^{quasi}_{min} and $D^{b-bistar}_{min}$.

It has been shown that [14]

$$D^{star}_{max} = \binom{n}{2}.$$

D^{linear}_{max} is unknown but $D^t_{max,P}$, the maximum value of D over the all the planar (P) linear arrangements of a tree t, has been investigated. A linear arrangement is planar if there are no edge crossings [25]. Many real spatial networks in two dimensions are planar or quasi planar [3]. In one dimension, the concept of a planar linear arrangement has applications in areas like circuit layout [26] or syntax [27].

It has been shown that [14]

$$D^{star}_{max} = D^{linear}_{max,P} = \binom{n}{2}.$$

Notice that edge crossings are impossible in a star tree [14] and hence $D^{star}_{max} = D^{star}_{max,P}$. Here we will calculate D^{linear}_{max} as well as D^{quasi}_{max} and $D^{b-bistar}_{max}$.

(ii) D^t_{min} and D^t_{max} in classes of trees (comprising more than one distinct unlabelled tree but not all distinct labelled trees). Two classes are selected for their theoretical interest: bistar trees (for the reasons explained above) and caterpillar trees (*cat*). Caterpillar trees is the class of trees such that when all the leaves are removed a linear tree is left [28]. Caterpillar trees are relevant for being a generalization of linear trees and bistar trees of enough simplicity that simple formulae for D^t_{min} can be obtained [29]. For each relevant class, we aim to express D^t_{min} and D^t_{max} as function of n and additional parameters of the networks extracted from vertex degrees: e.g., k_1, the largest degree, or $\langle k^2 \rangle$, the second moment of degree about zero.

(iii) The variation of D^t_{min} and D^t_{max} over all distinct unlabelled trees of n vertices. The problem is motivated by research on D^t as a function of n [2, 30, 31, 32]. It is well-known that any tree t of n vertices satisfies [20]

$$D^{linear}_{min} \leq D^t_{min} \leq D^{star}_{min} \leq D_{rla}. \tag{4}$$

The part $D^t_{min} \leq D^{star}_{min}$ is due to Iordanskii [13] although rediscovered later [20]. An inequality equivalent to equation 4 for D^t_{max} is not forthcoming but it has been shown that any tree t of n vertices satisfies [14]

$$D^t_{max,P} \leq D^{linear}_{max,P} = D^{star}_{max}.$$

Here we will show that any tree t of n vertices also satisfies

$$D_{rla} \leq D^{star}_{max} \leq D^t_{max} \leq D^{b-bistar}_{max}. \tag{5}$$

Table 1 summarizes all the existing results and the new results that are presented in this article for Problems i and ii.

The remainder of the article is organized as follows. Section 2 investigates D^t_{max}. Section 3 investigates D^t_{min}. Applying findings from the preceding sections, Section 4 investigates the limits of the variation of the optimality scores Δ^t and Γ^t while Section 5 investigates those of D^t_z. Finally, Section 6 reviews all our findings and suggests future research problems.

2. The maximum value of D

Here we investigate D^t_{max} in linear trees and bistar trees as well as the limits of the variation of D^t_{max} over all trees of n vertices.

2.1. Linear trees

A linear tree is a tree whose vertices are linked as a chain, i.e., a tree with arcs of the form $\{v_1, v_2\}, \{v_2, v_3\}, \ldots, \{v_{n-1}, v_n\}$. We will show that the maximum value of D^{linear} over the $n!$ linear arrangements is

$$\begin{aligned} D^{linear}_{max} &= \frac{1}{2}(n^2 - n \bmod 2) - 1 \\ &= \left\lfloor \frac{n^2}{2} \right\rfloor - 1 \end{aligned} \tag{6}$$

applying a result by [33]. Given a set $A_n = \{a_1, a_2, \ldots, a_n\}$, where $a_1 < a_2 < \ldots < a_n$, [33] shows how to calculate a permutation $\alpha = \alpha_1\alpha_2\ldots\alpha_n$ such that for certain functions f,

$$D_f(\alpha) = \sum_{i=1}^{n-1} f(|\alpha_i - \alpha_{i+1}|)$$

is maximized. D^{linear}_{max} is a particular case with $A_n = \{1, 2, \ldots, n\}$ and f the identity function (id).

Table 1. D^t_{min} and D^t_{max}, the minimum and the maximum value of D, the sum of edge lengths of a tree t. n is the number of vertices of the tree and k_1 is the largest degree, q is the number of vertices of odd degree, $q' = k_1 \bmod 2 + (n - k_1) \bmod 2$ and $\phi = ((n+2)^2 \bmod 8)$. For D^t_{min} we provide at least two formulae: one based on the floor or ceil function and the other based on mod (the exception are linear trees due to the simplicity of their formula). Formulae without a reference attached are new to our knowledge.

t	D^t_{min}	D^t_{max}
caterpillar	$n - 1 + \sum_{i=1}^{n} \left\lfloor \frac{1}{4}(k_i - 1)^2 \right\rfloor$ [29] $\sum_{i=1}^{n} \left\lfloor \frac{1}{4}(k_i + 1)^2 \right\rfloor - (n-1)$ $\frac{1}{4}\left(n\left\langle k^2 \right\rangle + q\right)$	
linear	$n - 1$ [12]	$\left\lfloor \frac{n^2}{2} \right\rfloor - 1$ $\frac{1}{2}(n^2 - n \bmod 2) - 1$
bistar	$\left\lfloor \frac{1}{4}(k_1+1)^2 \right\rfloor + \left\lfloor \frac{1}{4}(n - k_1 + 1)^2 \right\rfloor - 1$ $\frac{1}{2}k_1(k_1 - n) + \frac{1}{4}\left[n(n+2) + q'\right] - 1$	$k_1(n - k_1) + \frac{n}{2}(n-3) + 1$
balanced bistar	$\left\lfloor \frac{1}{8}(n+2)^2 \right\rfloor - 1$ $\frac{1}{8}(n^2 + 4n - 4 - \phi)$	$\frac{1}{4}\left(3(n-1)^2 + 1 - n \bmod 2\right)$
quasistar	$\left\lfloor \frac{1}{4}(n-1)^2 \right\rfloor + 1$ $\frac{1}{4}[n(n-2) + n \bmod 2] + 1$	$\frac{1}{2}(n+3)(n-2)$
star	$\left\lfloor \frac{1}{4}n^2 \right\rfloor$ [13] $\frac{1}{4}(n^2 - n \bmod 2)$ [20]	$\binom{n}{2}$ [14]

Under these assumptions, for each permutation α of the form noted above, we can construct a linear arrangement whose cost is $D_{id}(\alpha)$. This is the arrangement where the i-th vertex of the chain (v_i) is assigned the position α_i in the linear arrangement. Thus, the length of each arc of the form $\{v_i, v_{i+1}\}$ is $|\alpha_i - \alpha_{i+1}|$, and the total sum of lengths is $\sum_{i=1}^{n-1} |\alpha_i - \alpha_{i+1}|$, which equals $D_{id}(\alpha)$. This correspondence between permutations and linear arrangements is trivially bijective, as one can go from linear arrangements to permutations following the inverse process.

We restate Theorem 1 in [33] under our particular conditions as follows:

Theorem 1 (Chao and Liang, [33])**.** *A permutation of $\{1, 2, ..., n\}$ is maximum if it maximizes*

$$D(\alpha) = \sum_{i=1}^{n-1} |\alpha_i - \alpha_{i+1}|.$$

If $n = 2c$, then the maximum permutations with $\alpha_1 > \alpha_n$ are those satisfying the following three conditions:

(i) $\alpha_1 = c + 1$, $\alpha_n = c$

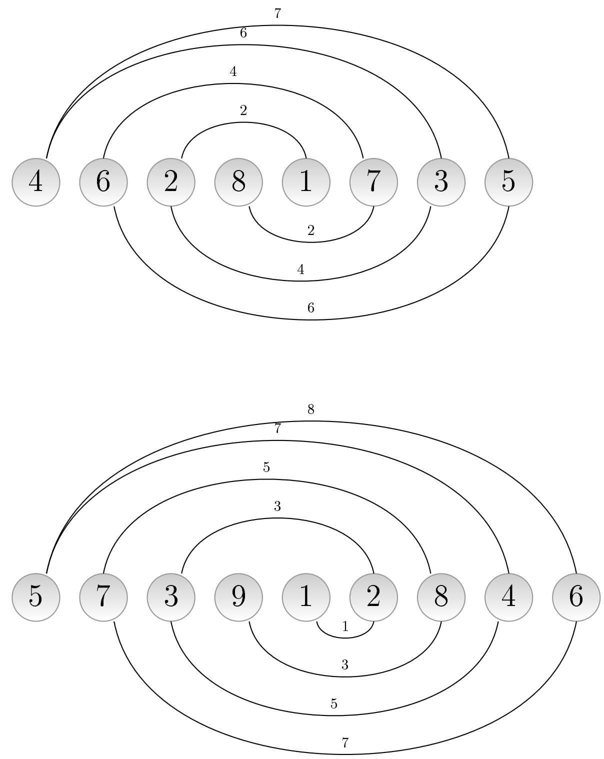


Figure 3. Maximum linear arragements of linear trees with n vertices. Vertex labels indicate the position of each vertex in the degree sequence. Edge labels indicate edge lengths. (a) $n = 8$ and $D^t = D^{linear}_{max} = 31$, generated by the permutation $\alpha = 5, 3, 7, 1, 8, 2, 6, 4$. (b) $n = 9$ and $D^t = D^{linear}_{max} = 39$, generated by the permutation $\alpha = 5, 6, 3, 8, 1, 9, 2, 7, 4$.

(ii) $\alpha_2\alpha_4 \cdots \alpha_{2c-2}$ *is a permutation of* $\{1, 2, \dots, c-1\}$

(iii) $\alpha_3\alpha_5 \cdots \alpha_{2c-1}$ *is a permutation of* $\{c+2, c+3, \dots, 2c\}$

If $n = 2c + 1$*, then the maximum permutations with* $\alpha_1 > \alpha_n$ *are those satisfying the following three conditions:*

(i) $\alpha_1 = c + 1$*,* $\alpha_n = c$

(ii) $\alpha_2\alpha_4 \cdots \alpha_{2c}$ *is a permutation of* $\{c+2, c+3, \dots, 2c+1\}$

(iii) $\alpha_3\alpha_5 \cdots \alpha_{2c-1}$ *is a permutation of* $\{1, 2, \dots, c-1\}$

or the following three conditions:

(i) $\alpha_1 = c + 2$*,* $\alpha_n = c + 1$

(ii) $\alpha_2\alpha_4 \cdots \alpha_{2c}$ *is a permutation of* $\{1, 2, \dots, c\}$

(iii) $\alpha_3\alpha_5 \cdots \alpha_{2c-1}$ *is a permutation of* $\{c+3, c+4, \dots, 2c+1\}$

The maximum permutations with $\alpha_1 < \alpha_n$ *are the reverse permutations of those specified above.*

Figure 3 shows arrangements that correspond to permutations that satisfy the conditions and, therefore, have maximum sum of lengths. By adding each of the lengths, it can be seen that the value of the sum for such an arrangement, when n is even (Figure 3 (a)), is

$$\begin{aligned}
D_{max}^{linear} &= n-1+2\times(n-2)+2\times(n-4)+\ldots+2\times 4+2\times 2 \\
&= n-1+2\times[(n-2)+(n-4)+\ldots+4+2] \\
&= n-1+2\times\left[\frac{n-2}{2}\times\frac{2+n-2}{2}\right] \\
&= n-1+\frac{n^2-2n}{2} \\
&= \frac{n^2-2}{2}. \qquad (7)
\end{aligned}$$

When n is odd (Figure 3 (b)), reasoning analogously we reach the sum

$$\begin{aligned}
D_{max}^{linear} &= n-1+2\times(n-2)+2\times(n-4)+\ldots+2\times 3+1 \\
&= n-1+2\times[(n-2)+(n-4)+\ldots+5+3]+1 \\
&= n+2\times\left[\frac{n-3}{2}\times\frac{3+n-2}{2}\right] \\
&= n+\frac{n^2-2n-3}{2} \\
&= \frac{n^2-3}{2}. \qquad (8)
\end{aligned}$$

Finally, equations 7 and 8 can be unified as equation 6.

2.2. Bistar trees

Hereafter we assume that a vertex is labelled with its position in the degree sequence, namely the non-increasing sequence of vertex degrees. Then k_i is the degree of the vertex with the i-th largest degree. A bistar tree is a generalization of trees of high theoretical interest: star trees [13, 20] and quasistar trees [23, 24]. If $k_1 = n-1$ (hence $k_2 = 1$) then we have a star tree. If $k_1 = n-2$ (hence $k_2 = 2$) then we have a quasistar tree (figures 1 (b-d)). Since a bistar tree consists of two joined stars, one may think that a bistar tree has three parameters, n, k_1 and k_2. However, n and k_1 suffice, as we will see next.

A bistar tree with $n \geq 2$ vertices satisfies the following properties:

(i) It has at most two internal vertices, more precisely $2-\delta_{k_1,1}-\delta_{k_2,1}$ internal vertices, where δ is the Kronecker delta function. $\delta_{k_2,1} = 1$ when the tree is a star.

(ii) It has $n-2+\delta_{k_1,1}+\delta_{k_2,1}$ leaves.

(iii) Then $k_i = 1$ for $3-\delta_{k_1,1}-\delta_{k_2,1} \leq i \leq n$.

(iv) $$k_2 = n - k_1 \qquad (9)$$
because the sum of vertex degrees must satisfy
$$k_1 + k_2 + n - 2 = 2(n-1)$$

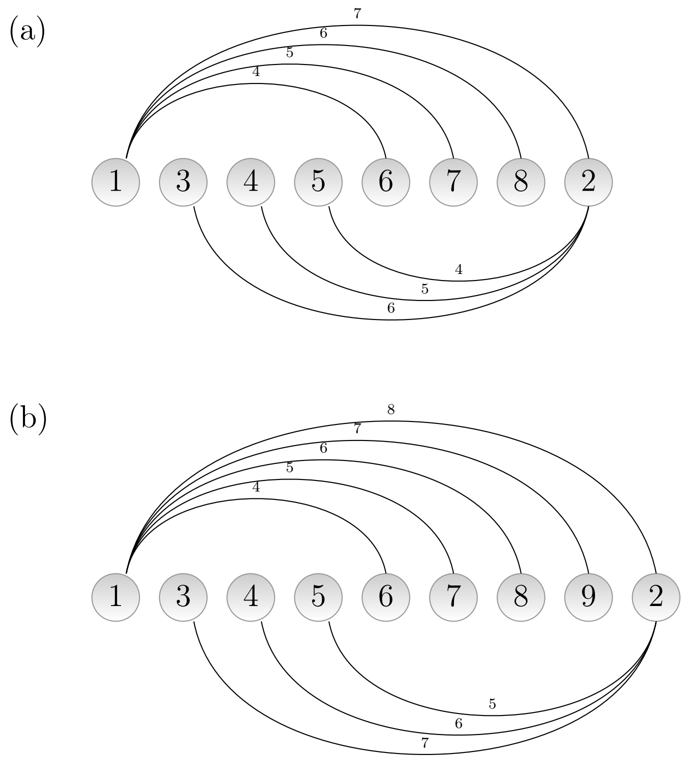


Figure 4. Extreme linear arrangements of balanced bistar trees of n vertices. Vertex labels indicate the position of each vertex in the degree sequence. Edge labels indicate edge lengths. (a) $n = 8$ and $D^t = D_{max}^{b-bistar} = 37$. (b) $n = 9$ and $D^t = D_{max}^{b-bistar} = 48$.

by the handshaking lemma.

(v)

$$\begin{aligned} \langle k^2 \rangle &= \frac{1}{n}(k_1^2 + (n-k_1)^2 + n - 2) \\ &= \frac{2}{n}(k_1(k_1-n) - 1) + n + 1 \end{aligned} \tag{10}$$

(vi)
$$k_1 \geq \left\lceil \frac{n}{2} \right\rceil . \tag{11}$$

Combining equation 9 with the condition $k_1 \geq k_2$, one obtains

$$k_1 \geq \frac{n}{2},$$

which knowing that k_1 is an integer gives equation 11.

Our definition of a star tree with two parameters, n and k_1, is equivalent to other two-parameter definitions. [21] defines a bistar with two parameters n_1 and n_2. The bistar is formed by taking a graph with a single edge and two vertices and adding n_1 edges at one end of the edge and n_2 edges at the other end. Ours is then $n = 2 + n_1 + n_2$ and $k_1 = \max(n_1, n_2) + 1$. [22] defines a bistar with two parameters n'_1 and n'_2. The bistar is formed by adding an edge between the hubs of two stars of n'_1 and n'_2 vertices respectively [22]. Ours is then $n = n'_1 + n'_2$ and $k_1 = \max(n'_1, n'_2)$. The term bistar tree has also been used to refer to a tree with only one inner edge or a tree of diameter three, where the diameter is the length of the longest shortest path in edges [34]. This is not exactly our definition of bistar because it excludes star trees and implies $n \geq 4$. In our definition, a bistar tree has at most one inner edge and diameter at most 3 and is then valid for $n < 4$.

We introduce a bistar tree of great theoretical importance to calculate the maximum of D^t_{max} over all trees of same size: the balanced bistar tree (figure 4). That tree is a bistar tree with

$$k_1 = \left\lceil \frac{n}{2} \right\rceil. \tag{12}$$

The latter implies that $k_2 = \left\lfloor \frac{n}{2} \right\rfloor$ thanks to equation 9. The term balanced comes from the fact a balanced bistar tree is a bistar tree where the difference $k_1 - k_2$ is minimized. Thanks to equation 9, one has that

$$k_1 - k_2 = 2k_1 - n.$$

The fact that $k_1 \geq k_2$, gives that the difference is minimized when k_1 satisfies equation 12.

2.2.1. D^{bistar}_{max}, D^t_{max} in bistar trees We define $\gamma(i)$ as the set of adjacent vertices of i [35], also termed the set of 1st neighbours or nearest neighbours of i [36]. We define an extreme linear arrangement of a bistar tree as an ordering of the vertices following the one of the following templates:

$$1, \gamma(2) \setminus \{1\}, \gamma(1) \setminus \{2\}, 2$$

as in figure 4, or its symmetric, i.e.

$$2, \gamma(1) \setminus \{2\}, \gamma(2) \setminus \{1\}, 1$$

The following lemma indicates how to arrange a single vertex and its attached vertices so as to maximize its sum of edge lengths.

Lemma 1. *Suppose that D^g_i is the sum of the lengths of the edges attached to the i-th vertex of a graph g of n vertices. Then $D^g_{i,max}$, the maximum value of D^g_i over the $n!$ linear arrangements of the whole graph is*

$$D^g_{i,max} = \frac{1}{2} k_i (2n - k_i - 1) \tag{13}$$

and is achieved when vertex i is placed at one of the ends of the linear arrangement and its adjacent vertices at the other end.

Proof. When the i-th vertex is placed at one of the ends of the linear arrangement and its adjacent vertices as far as possible (consecutively at the other end),

$$D^g_i = \sum_{j=1}^{k_i} (n - j),$$

which gives equation 13. If the i-th vertex is not placed at one of the ends but its neighbours are still placed as far as possible, D_i^g cannot exceed $D_{i,max}^g$. A detailed argument follows.

We define h as the position of vertex i in the linear arrangement ($1 \leq h \leq n$), k_i^- as the number of neighbours of i placed before i and k_i^+ as the number of neighbours of i placed after i. In such a linear arrangement, the maximum value of D_i^g, i.e. $D_{i,max,k_i^-,k_i^+}^g$, is achieved placing the k_i^- neighbours at the beginning of the linear arrangement and the k_i^+ neighbours at the end of the linear arrangement, producing

$$D_{i,max,k_i^-,k_i^+}^g = \sum_{j=1}^{k_i^-}(h-j) + \sum_{j=1}^{k_i^+}(n-j+1-h).$$

We will show that $D_{i,max,h,k_i^-,k_i^+}^g \leq D_{i,max}^g = D_{i,max,1,0,k_i}^g = D_{i,max,n,k_i,0}^g$, i.e.

$$\sum_{j=1}^{k_i^-}(h-j) + \sum_{j=1}^{k_i^+}(n-j+1-h) \leq \sum_{j=1}^{k_i}(n-j),$$

that is equivalent to

$$\sum_{j=1}^{k_i-k_i^+}(h-j) + (1-h)k_i^+ \leq \sum_{j=k_i^+ +1}^{k_i}(n-j).$$

Rearranging the terms and calculating certain summations the inequality becomes

$$0 \leq (k_i - k_i^+)(n-h) + \sum_{j=1}^{k_i-k_i^+} j - \sum_{j=k_i^+ +1}^{k_i} j + (h-1)k_i^+.$$

Calculating the remaining summations one obtains, after some routine calculations,

$$0 \leq (k_i - k_i^+)(n-h-k_i^+) + (h-1)k_i^+,$$

which is trivially true because $k_i^+ \leq k_i$, $1 \leq h$ and $k_i^+ \leq n-h$ by definition. □

The previous lemma generalizes a previous result on D_{max}^{star}, that is achieved when the hub of the star is located at one of the ends of the linear arrangement [37] (figure 7 (a)). In a star tree t, D^t is determined by the sum of edge lengths of the hub vertex.

The following lemma indicates that an extreme linear arrangement of a bistar is actually a maximum linear arrangement.

Lemma 2. *In a bistar tree t of n vertices and maximum degree k_1, D_{max}^t is*

$$D_{max}^{bistar} = k_1(n-k_1) + \frac{n}{2}(n-3) + 1 \tag{14}$$

and a extreme linear arrangement of t is actually a maximum linear arrangement.

Proof. A bistar tree can be seen as two star trees joined by a common edge. Then D^t can be decomposed as

$$D^t = D_1^t + D_2^t - d_{12}^t, \tag{15}$$

where D_i^t is the sum of the lengths of edges attached to the vertex with the i-th largest degree and d_{12}^t is the length of the edge joining the two vertices with the largest degrees. To maximize D^t following equation 15, one has to maximize D_1^t and D_2^t. By lemma 1, D_1^t is maximized placing vertex 1 at one end and its neighbours at the other end. Since D_2^t also must be maximized then, by the same lemma, vertex 2 has to be placed at the opposite end (otherwise $D_1^t < D_{1,max}^t$ or $D_2^t < D_{2,max}^t$), which gives $d_{12}^t = n-1$. Such a linear arrangement is an extreme linear arrangement of a bistar tree and equation 15 gives

$$\begin{aligned} D_{max}^{bistar} &= D_{1,max}^t + D_{2,max}^t - (n-1) \\ &= \frac{1}{2}k_1(2n-k_1-1) + \frac{1}{2}(n-k_1)(n+k_1-1) - (n-1). \end{aligned}$$

Equation 14 is recovered after some algebra. □

Thanks to the preceding work, formulae of D_{max}^t for specific bistar trees follow easily.

Corollary 1.

$$\begin{aligned} D_{max}^{b-bistar} &= \frac{1}{4}\left(3(n-1)^2 + 1 - n \bmod 2\right) \qquad (16) \\ D_{max}^{quasi} &= \frac{1}{2}(n+3)(n-2) \\ D_{max}^{star} &= \binom{n}{2}. \end{aligned}$$

Proof. $D_{max}^{b-bistar}$ is obtained applying $k_1 = \lceil n/2 \rceil$ to equation 14 (Theorem 2). When n is odd, $k_1 = (n+1)/2$ and then equation 14 gives

$$D_{max}^{b-bistar} = \frac{3}{4}(n-1)^2.$$

When n is even, $k_1 = n/2$, one obtains

$$D_{max}^{b-bistar} = \frac{1}{4}(3(n-1)^2 + 1).$$

Therefore, for any n, $D_{max}^{b-bistar}$ follows equation 16. Similarly, D_{max}^{quasi} is obtained with $k_1 = n-2$ and D_{max}^{star} is obtained with $k_1 = n-1$ after some mechanical work. □

2.3. The maximum D^t_{max}

In a graph g of n vertices and m edges, an obvious upper bound of D^t_{max} is [17]

$$D^g_{upper,naive} = m(n-1),$$

where $n-1$ is the maximum length of an edge. *A priori*, $n-d$ edges of length d can be formed. Taking m lengths as long as possible, the analog of Petit's edge method (EM) for the maximum linear arrangement problem [38], gave another upper bound of D^t_{max} [17] that is

$$\begin{aligned} D^g_{upper,EM} = (m - F(d_*))(d_* - 1) \\ + \frac{1}{6}(n-d_*)(n^2 + (n+3)d_* - 2d_*^2 - 1), \end{aligned}$$

where

$$\begin{aligned} F(d_0) &= \frac{1}{2}(n-d_0)(n-d_0+1) \\ d_* &= \left\lceil n + \frac{1}{2} - \frac{1}{2}\sqrt{8m+1} \right\rceil . \end{aligned}$$

Figure 5 shows that, when $m = n-1$ as in a tree, the naive upper bound, i.e. $(n-1)^2$, beats the edge method upper bound for sufficiently large n. This is likely to be due to the tree constraint (acyclicity and connectedness). Interestingly, the naive upper bound is close to the true maximum of D^t_{max}, that is achieved by a maximum linear arrangement of a balanced bistar tree as we will show next.

Theorem 2 (Maximum D^t_{max})**.** *For any tree t of n vertices,*

$$D^t_{max} \le D^{b-bistar}_{max} = \frac{1}{4}\left(3(n-1)^2 + 1 - n \bmod 2\right).$$

Proof. Let τ be the set of all unlabelled trees of n vertices. Let υ be the set of labelled trees of n vertices, i.e., the set of trees of n vertices where each vertex has been assigned a unique number in $\{1, 2, \ldots, n\}$ that indicates its position in the linear arrangement. Given an unlabelled tree $t \in \tau$, choosing a linear arrangement for it (by assigning a linear order to its vertices) results into one of the trees in υ. Thus, maximizing the value of D^t across the $n!$ possible linear arrangements of each unlabelled tree in τ reduces to maximizing the value of D in υ.‡

Let φ_1 be the set of directed rooted trees obtained by rooting each tree in υ at its vertex 1, the 1st vertex in the linear arrangement, and directing all edges to point away from the root. Trivially, this mapping between φ_1 and υ is bijective (as said orientation is unique) and it preserves the sum of edge lengths. Therefore, if we find a directed tree with maximum sum of arc lengths in φ_1, its underlying undirected tree will have the maximum sum of edge lengths in υ.

‡ Note that the mapping from linear arrangements to trees in υ is not bijective (different linear arrangements can result into the same labelled tree, e.g. all the linear arrangements of a star tree where the central vertex's position is kept constant) but this is not relevant for this proof, as it does not affect D.

We will show that the directed tree with arcs $n \to 2, n \to 3, \ldots, n \to \lfloor \frac{n}{2} \rfloor, 1 \to \lfloor \frac{n}{2} \rfloor + 1, \ldots, 1 \to n-1, 1 \to n$, whose underlying undirected tree is a balanced bistar tree as in figure 4, maximizes the sum of arc lengths in φ_1. To see this, we use the property of directed trees that every vertex has exactly one incoming arc, except for the root which has none. Thus, for any tree of φ_1, we can write its arcs as $A_2, A_3, \ldots, A_n$ such that A_i is the arc going into vertex i. Now, if we consider each arc individually, we can say that

- The length of the arc A_i, for $2 \leq i \leq \lfloor \frac{n}{2} \rfloor$, is at most $n - i$, as n is the farthest possible vertex from vertex i. That is, the arc $n \to i$ is the longest possible arc to vertex i.
- The length of the arc A_i, for $\lfloor \frac{n}{2} \rfloor < i \leq n$, is at most $i - 1$, as 1 is the farthest possible vertex from vertex i. That is, the arc $1 \to i$ is the longest possible arc to vertex i.

The directed tree mentioned above has exactly the longest arcs going into each vertex i, for $2 \leq i \leq n$. Thus, it maximizes the sum of arc lengths in φ_1, as it maximizes each individual length $A_2, \ldots, A_n$. Therefore, the underlying balanced bistar tree has the maximum sum of edge lengths in v, proving the theorem. Note: the ordering of the vertices implied by such directed tree corresponds to an extreme linear arrangement as described above. □

The problem of maximizing D^t_{max} is equivalent to finding the maximum spanning tree on a complete graph where the weight of every edge is the distance between the vertices that form it in the linear arrangement (Appendix A). The argument provides an alternative way to demonstrate Theorem 2.

2.4. The minimum D^t_{max}

The following theorem states that for any tree with n vertices we can always find an arrangement with total length at least $\frac{n(n-1)}{2}$.

Theorem 3 (Minimum D^t_{max})**.** *For any tree t of n vertices,*

$$D^t_{max} \geq D^{star}_{max} = \binom{n}{2}.$$

Proof. Let us consider a tree t of n vertices and define, $\gamma(v,i)$, the set of vertices that are at topological distance i from vertex v in t, with $\gamma(v,0) = \{v\}$. Equivalently, $\gamma(v,i)$ is the set of i-th neighbours of i. $\gamma(v,1)$ is the set of vertices adjacent to i [35]. For instance, in a star tree of n vertices where the hub is vertex 1 and $v \neq 1$ is some leaf, $\gamma(v,1) = \{1\}$ and $\gamma(v,2) = \{2,3,...,\} \setminus \{v\}$.

A linear arrangement that gives sum at least D^{star}_{max} follows the template defined by the sequence

$$\gamma(v,0), \gamma(v,2), \gamma(v,4), \cdots, \gamma(v,3), \gamma(v,1) \tag{17}$$

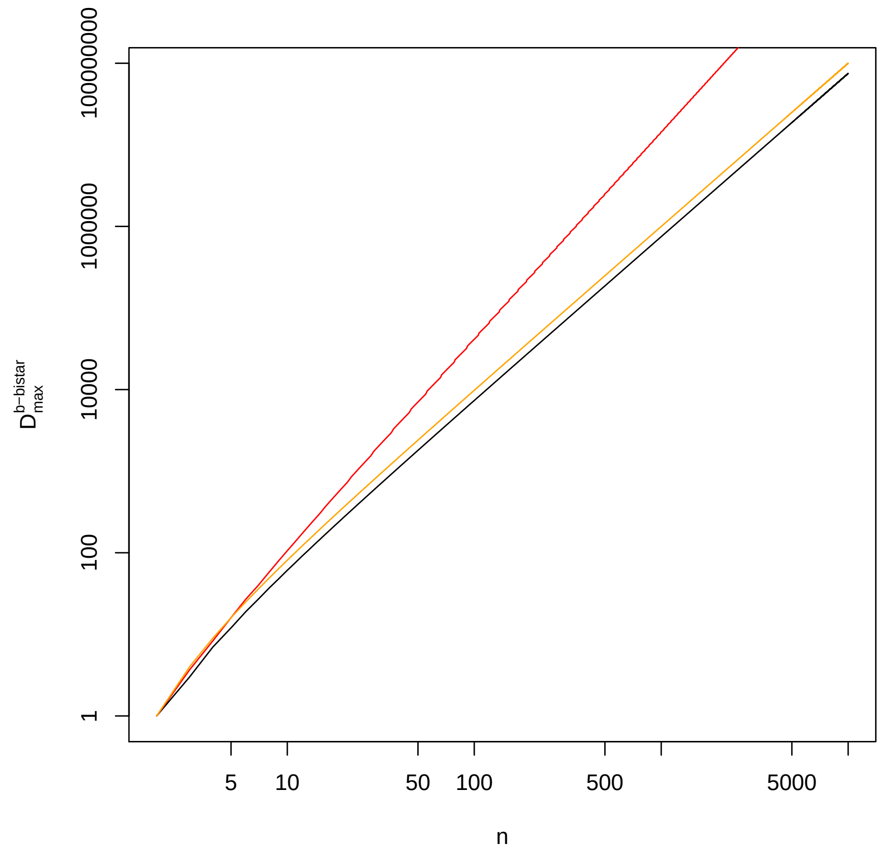


Figure 5. $D_{max}^{b-bistar}$, the true maximum of D_{max}^{t}, as a function of n (black). The predictions of the naive upper bound (orange) and the edge method upper bound (red) are also shown.

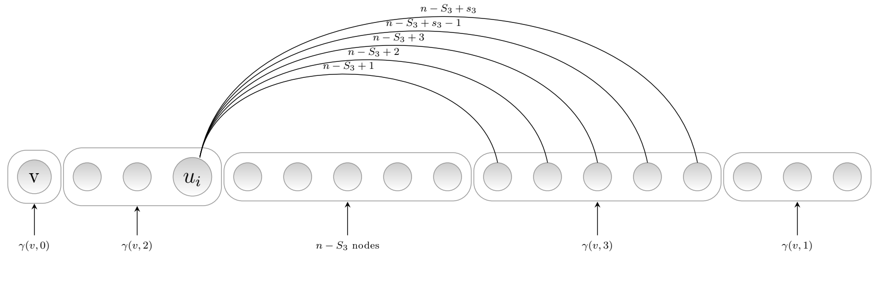


Figure 6. The edges between $\gamma(v, 2)$ and $\gamma(v, 3)$ when they are formed exclusively with u_i, the vertex of $\gamma(v, 2)$ that is the closest to $\gamma(v, 3)$.

This is not a proper arrangement because the $\gamma(v,i)$ is a set and its elements are not ordered. We can get a proper arrangement by ordering the vertices in every set $\gamma(v,i)$ in any arbitrary way.

Let s_i be $|\gamma(v,i)|$ and $S_i = \sum_{j=0}^{i} s_j$. We define V_i as the set of vertices reached up to topological distance i, i.e.

$$V_i = \cup_{j=0}^{i}\gamma(v,j).$$

Hence $S_i = |V_i|$.

Let us use induction on the topological distance i.

Induction hypothesis. The sum of the lengths of the edges formed by vertices in V_i is at least

$$\sum_{j=1}^{S_i-1}(n-j). \tag{18}$$

Base case. For $i=0$, $S_i=1$ and the sum of edge lengths must be zero trivially.

Induction step. Note that the number of vertices between $\gamma(v,i)$ and $\gamma(v,i+1)$, that is $s_{i+2}+s_{i+3}+\ldots$, is $n-S_{i+1}$ (figure 6). According to the template of linear arrangement in equation 17, the vertices in $\gamma(v,i+1)$ are the farthest away from those of $\gamma(v,i)$ among vertices with topological distance $i+1$ or more. Let B the sum of lengths of the s_{i+1} edges from $\gamma(v,i)$ to $\gamma(v,i+1)$. Suppose that these edges start from u_i, the vertex in $\gamma(v,i)$ nearer to $\gamma(v,i+1)$ in the linear arrangement, which implies the vertices in $\gamma(v,i)\setminus\{u_i\}$ must be leaves (figure 6). Then

$$\begin{aligned} B &= \sum_{j'=1}^{s_{i+1}}(n-S_{i+1}+j') \\ &= \sum_{j=S_{i+1}-s_{i+1}}^{S_{i+1}-1}(n-j) \\ &= \sum_{j=S_i}^{S_{i+1}-1}(n-j). \end{aligned}$$

If edges from $\gamma(v,i)$ to $\gamma(v,i+1)$ involved any vertex in $\gamma(v,i)\setminus\{u_i\}$, then B would increase. Thus, thanks to the induction hypothesis, the sum of the costs of the edges from the vertices in V_{i+1} is at least

$$\begin{aligned} \sum_{j=1}^{S_i-1}(n-j)+B &= \sum_{j=1}^{S_i-1}(n-j)+\sum_{j=S_i}^{S_{i+1}-1}(n-j) \\ &= \sum_{j=1}^{S_{i+1}-1}(n-j) \end{aligned}$$

as expected.

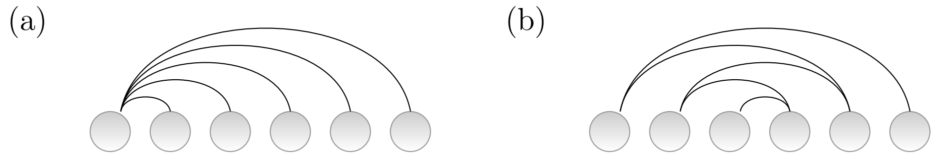


Figure 7. Linear arrangements of trees of n vertices where $n = 6$ and $D^t = \binom{n}{2} = 15$. (a) Linear arrangement of a star tree that is both a maximum linear arrangement and a maximum planar linear arrangement ($D^t = D^t_{max,P} = D^t_{max}$). (b) Linear arrangement of a linear tree that is a maximum planar linear arrangement but not a maximum linear arrangement ($D^t = D^t_{max,P} < D^t_{max} = 17$; recall Table 1).

Let $\zeta(v)$ be the maximal topological distance to v in some tree, then $S_{\zeta(u)} = n$ and equation 18 gives

$$D^t_{max} \geq \sum_{j=1}^{n-1} j = \binom{n}{2}.$$

□

The previous theorem indicates that D^t_{max} is at least its value for star trees (Figure 7). However, it is well-known that D^{star}_{max} can also be achieved by a linear tree arranged as in Figure 7 [14]. That arrangement follows from applying the template of arrangement in equation 17 with one of the leaves as the initial vertex.

2.5. The relationship with D_{rla}

By definition of average and maximum, $D_{rla} \leq D^t_{max}$, and particularizing this for a star tree, $D_{rla} \leq D^{star}_{max}$. Putting this together with Theorem 3, we have that $D_{rla} \leq D^{star}_{max} \leq D^t_{max}$.

3. The minimum value of D

D^t_{min} can be calculated with rather complex algorithms for any tree t [12, 10, 11]. Algorithms to calculate D^t_{min}, the minimum value of D over the all the linear arrangements of a tree t satisfying a certain constraint, are also available but less known. See [39, 40] for planarity (no edge crossings) and [41, 30] for projectivity. Here we are interested in compact formulae for D^t_{min} for certain classes of trees or general lower bounds.

[38] reviews various techniques to obtain lower bounds of D^t. In a network with m edges, the edge method consists of picking the m shortest edges noting that there can be at most $n - d$ edges of length d, for $1 \leq d \leq n - 1$. In a tree, this methods trivially gives $D_{min} \geq D^{linear}_{min} = n - 1$. The next theorem presents a lower bound of D^t_{min} that depends exclusively on the degree sequence and that is obtained with the degree method. A similar application of the degree method can be found in [14].

Theorem 4. *For any tree t of n vertices,*

$$D^t_{min} \geq D^t_0 = \frac{1}{4}\left(\frac{n}{2}\left\langle k^2 \right\rangle + 2(n-1) + \frac{1}{2}q\right)$$

where

$$q = \sum_{i=1}^{n} k_i \bmod 2$$

is the number of vertices of odd degree.

Proof. Let D^t_i be the sum of the length of the edges attached to the i-th vertex of t. The degree method is based on a star tree decomposition of D in a network, whereby [38]

$$D^t = \frac{1}{2}\sum_{i=1}^{n} D^t_i. \tag{19}$$

$D^t_{i,min}$ is a lower bound of D^t_i defined as [38]

$$D^t_{i,min} = 2\sum_{j=1}^{k_i/2} j = \frac{1}{2}\left(\frac{k_i^2}{2} + k_i\right)$$

if k_i is even, and

$$D^t_{i,min} = \frac{1}{2}(k_i+1) + 2\sum_{j=1}^{(k_i-1)/2} j = \frac{1}{2}\left(\frac{k_i^2}{2} + k_i + \frac{1}{2}\right)$$

if k_i is odd. Combining both results, one obtains

$$D^t_{i,min} = \frac{1}{2}\left(\frac{k_i^2}{2} + k_i + \frac{1}{2}k_i \bmod 2\right).$$

Inserting the last result into equation 19, one obtains

$$\begin{aligned}
D^t_{min} &\geq \frac{1}{2}\sum_{i=1}^{n} D_{i,min} \\
&= \frac{1}{4}\sum_{i=1}^{n}\left(\frac{k_i^2}{2} + k_i + \frac{1}{2}k_i \bmod 2\right) \\
&= \frac{1}{4}\left(\frac{n}{2}\left\langle k^2 \right\rangle + 2(n-1) + \frac{1}{2}q\right) \\
&= D^t_0.
\end{aligned}$$

□

In a linear tree, $\langle k^2 \rangle = 4 - 6/n$ [14] and $q = 2$ give $D_0 = n - 1$, matching D^{linear}_{min}. In contrast, for a star tree, $\langle k^2 \rangle = n - 1$ [14] and $q = n - 1 + n - 1 \bmod 2$ give

$$D_0 = \frac{1}{8}\left(n^2 + 4n - 5 + (n-1) \bmod 2\right) \tag{20}$$

while $D_{min}^{star} = \frac{1}{4}(n^2 - n \bmod 2)$ [13]. Asymptotically, D_0 deviates from the true minimum, D_{min}^{star}, by a factor of 1/2.

The following theorem is a formalization of the arguments of [29] that presents a lower bound of D_{min}^t that has no deviation if the tree is a caterpillar (*cat*), including then the particular cases of star trees and linear trees discussed above.

Theorem 5 (Horton [29])**.** *For any tree t of n vertices,*

$$D_{min}^t \geq D_{min}^{cat} \tag{21}$$

where D_{min}^{cat} is the value of D_{min}^t of a caterpillar tree with the same degree sequence as t. We have

$$D_{min}^{cat} = n - 1 + \sum_{i=1}^{n} a_i, \tag{22}$$

where

$$\begin{aligned} a_i &= \left\lfloor \frac{k_i}{2} \right\rfloor \left\lceil \frac{k_i - 2}{2} \right\rceil \\ &= \left\lfloor \frac{(k_i - 1)^2}{4} \right\rfloor \end{aligned} \tag{23}$$

and k_i is the degree of the i-th vertex.

Proof. In trees, the bipartite crossing number is [42]

$$bcr = D_{min} - n + 1 - \sum_{i=1}^{n} a_i. \tag{24}$$

$bcr \geq 0$ by definition and $bcr = 0$ if and only if the tree is a caterpillar tree [43]. Therefore, equation 24 becomes equation 21 with equality if and only of the tree is a caterpillar. A longer proof where a_i is expressed as in equation 23 is found in [29]. □

The following theorem introduces useful algebraic expressions for caterpillar trees, alternating floor with modulo operations.

Theorem 6.

$$D_{min}^{cat} = n - 1 + \sum_{i=1}^{n} \left\lfloor \frac{1}{4}(k_i - 1)^2 \right\rfloor \tag{25}$$

$$= \sum_{i=1}^{n} \left\lfloor \frac{1}{4}(k_i + 1)^2 \right\rfloor - (n - 1) \tag{26}$$

$$= \frac{1}{4}\left(n \left\langle k^2 \right\rangle + q\right). \tag{27}$$

Proof. Equation 25 is due to [29]. By Theorem 5 and Table 1,

$$\begin{aligned} \left\lfloor \frac{n-1}{2} \right\rfloor \left\lceil \frac{n-1}{2} - 1 \right\rceil &= D_{star}^{min} - (n-1) \\ &= \frac{n^2 - n \bmod 2}{4} - (n-1) \\ &= \left\lfloor \frac{1}{4} n^2 \right\rfloor - (n-1). \end{aligned}$$

The change of variable $k_i = n - 1$ gives

$$\begin{aligned} a_i &= \frac{(k_i+1)^2 - (k_i+1) \bmod 2}{4} - k_i \\ &= \left\lfloor \frac{1}{4}(k+1)^2 \right\rfloor - k_i. \end{aligned}$$

Plugging these results into equation 22 one obtains

$$\begin{aligned} D_{min}^{cat} &= n - 1 + \sum_{i=1}^{n} a_i \\ &= n - 1 + \sum_{i=1}^{n} \left(\left\lfloor \frac{1}{4}(k+1)^2 \right\rfloor - k_i \right) \\ &= n - 1 + \sum_{i=1}^{n} \left\lfloor \frac{1}{4}(k+1)^2 \right\rfloor - \sum_{i=1}^{n} k_i \\ &= \sum_{i=1}^{n} \left\lfloor \frac{1}{4}(k_i+1)^2 \right\rfloor - (n-1) \end{aligned}$$

and also

$$\begin{aligned} D_{min}^{cat} &= n - 1 + \sum_{i=1}^{n} a_i \\ &= n - 1 + \frac{1}{4}\sum_{i=1}^{n} \left[k_i^2 - 2k_i + 1 - (k_i+1) \bmod 2 \right] - \sum_{i=1}^{n} k_i \\ &= \frac{1}{4}\sum_{i=1}^{n} \left[k_i^2 + k_i \bmod 2 \right] + \frac{1}{2}\sum_{i=1}^{n} k_i - (n-1) \\ &= \frac{1}{4}\left(n \left\langle k^2 \right\rangle + q \right). \end{aligned}$$

□

It is easy to see that D_{min}^{cat} is a tighter lower bound of D_{min}^{t} than D_0^t. Thanks to equations 20 and 27, the condition $D_{min}^{cat} \geq D_0^t$ is equivalent to

$$\left\langle k^2 \right\rangle \geq 4\left(1 - \frac{1}{n}\right) - \frac{q}{n}.$$

Furthermore, this condition will be always satisfied provided that $n \geq 2$ because it holds even when $\langle k^2 \rangle$ takes its minimum value, namely [14]

$$\left\langle k^2 \right\rangle^{linear} = 4 - \frac{6}{2}.$$

The substitution by $\langle k^2 \rangle^{linear}$ in the condition above gives $q \geq 2$, which is trivially true for any tree such that $n \geq 2$ as any tree with such a number of vertices has at least two leaves.

The following corollary presents formulae of D_{min}^{t} for bistar trees and three instances: stars, quasistars and balanced star trees.

Corollary 2. *In any bistar tree, where k_1 is the largest degree,*

$$D_{min}^{bistar} = \left\lfloor \frac{1}{4}(k_1+1)^2 \right\rfloor + \left\lfloor \frac{1}{4}(n-k_1+1)^2 \right\rfloor - 1 \quad (28)$$

$$= \frac{1}{2}k_1(k_1-n) + \frac{1}{4}\left[n(n+2)+q'\right] - 1, \quad (29)$$

where

$$q' = k_1 \bmod 2 + (n-k_1) \bmod 2.$$

In addition,

$$D_{min}^{b-bistar} = \left\lfloor \frac{1}{8}(n+2)^2 \right\rfloor - 1 \quad (30)$$

$$= \frac{1}{8}(n^2+4n-4-\phi), \quad (31)$$

with

$$\phi = ((n+2)^2 \bmod 8),$$

and also

$$D_{min}^{quasi} = \left\lfloor \frac{1}{4}(n-1)^2 \right\rfloor + 1$$

$$= \frac{1}{4}[n(n-2) + n \bmod 2] + 1, \quad (32)$$

$$D_{min}^{star} = \left\lfloor \frac{1}{4}n^2 \right\rfloor$$

$$= \frac{1}{4}(n^2 - n \bmod 2). \quad (33)$$

Proof. As a bistar tree is a caterpillar tree, the application of equation 26 (Theorem 6) with $k_2 = n - k_1$ and $k_i = 1$ for $i \geq 3$, gives equation 28. Besides, the application of 27 (Theorem 6) with $\langle k^2 \rangle$ for a bistar tree (equation 10) produces equation 29 after some mechanical work.

As a balanced bistar tree is a bistar tree with $k_1 = \lceil n/2 \rceil$, equation 28 gives

$$\left\lfloor \frac{1}{4}\left(\left\lceil \frac{n}{2} \right\rceil + 1\right)^2 \right\rfloor + \left\lfloor \frac{1}{4}\left(\left\lfloor \frac{n}{2} \right\rfloor + 1\right)^2 \right\rfloor - 1$$

immediately. However, a much simpler modular formula will be obtained from equation 29. In this respect, notice that

$$k_1 = \frac{1}{2}(n + n \bmod 2),$$

and also that $0 \leq q' \leq 2$, in particular, $q' = 1$ if n is odd and

$$q' = 2\left(\frac{n}{2} \bmod 2\right)$$

if n is even. Then equation 29 produces

$$D_{min}^{b-bistar} = \begin{cases} \frac{1}{8}(n^2+4n-8+4(n/2 \bmod 2)) & n \text{ is even} \\ \frac{1}{8}(n^2+4n-5) & n \text{ is odd.} \end{cases}$$

or

$$D_{min}^{b-bistar} = \begin{cases} \frac{1}{8}(n^2+4n-8) & n \bmod 4 = 0 \\ \frac{1}{8}(n^2+4n-4) & n \bmod 4 = 2 \\ \frac{1}{8}(n^2+4n-5) & \text{otherwise.} \end{cases}$$

in expanded form. From this point, equation 31 follows immediately. Noting that

$$D_{min}^{b-bistar} = \frac{1}{8}[(n+2)^2 - 8 - (n+2)^2 \bmod 8]$$

and applying the definition of modulus, i.e.

$$(n+2)^2 \bmod 8 = (n+2)^2 - 8\left\lfloor \frac{(n+2)^2}{8} \right\rfloor,$$

one finally obtains 30.

A quasistar tree is a bistar tree where $k_1 = n-2$, which transforms equations 30 and 31 into equation 32 after some algebraic work. Similarly, a star tree is a bistar tree where $k_1 = n-1$, which transforms equations 30 and 31 into equation 33. Equation 33 has been derived through other means [20]. □

4. The maxima of optimality scores

Here we aim to investigate a couple of optimality scores: $\Delta^t = D^t - D^t_{min}$ [16] and $\Gamma^t = D^t/D^t_{min}$ [2, 15]. By definition of D^t_{min}, $\Delta^t \geq \Delta^t_{min} = 0$ and $\Gamma^t \geq \Gamma^t_{min} = 1$.

For a specific tree t of n vertices, the maximum value of Δ^t over all possible linear arrangements is

$$\Delta^t_{max} = D^t_{max} - D^t_{min}.$$

Similarly,

$$\Gamma^t_{max} = \frac{D^t_{max}}{D^t_{min}}.$$

Table 1 allows one to obtain formulae of Δ^t_{max} or Γ^t_{max} for specific trees. Figures 8 and 9 show the growth of Δ^t_{max} and Γ^t_{max} for specific trees. The star tree is actually a baseline because we will show that it minimizes Δ^t_{max} and Γ^t_{max}. In star trees, quasistar trees and balanced bistar trees, Γ^t_{max} converges to a constant (figure 9) because both D^t_{min} and D^t_{max} are quadratic functions of n (Table 1). In balanced bistar trees, the leading coefficients are 1/8 and 3/4, respectively, which gives

$$\lim_{n\to\infty} \Gamma^{b-bistar}_{max} = 6.$$

By similar arguments,

$$\lim_{n\to\infty} \Gamma^{star}_{max} = \lim_{n\to\infty} \Gamma^{quasi}_{max} = 2.$$

These limiting values are consistent with figure 9.

Here we aim to apply the results in the preceding sections to investigate an important question for research on these scores as a function of n [2, 16]: what are the minimum and the maximum value that Δ^t_{max} or Γ^t_{max} can attain over all trees of n vertices?

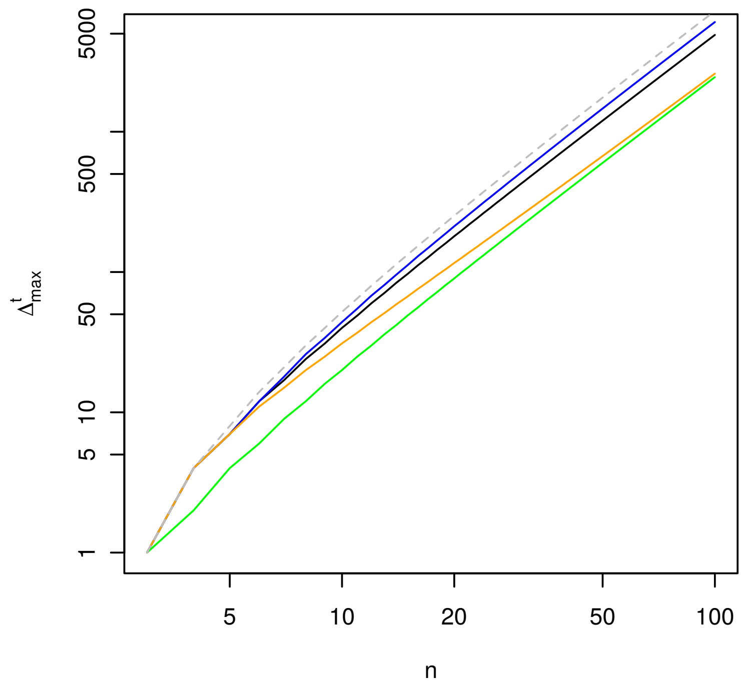


Figure 8. The scaling of Δ^t_{max} as a function of n, the number of vertices of the tree t for different trees: linear trees (black), quasistar trees (orange), star trees (green) and balanced bistar trees (blue). For reference, the upper bound $D^{b-bistar}_{max} - D^{linear}_{min}$ (dashed gray line) is also shown.

4.1. The minima of Δ^t_{max} and Γ^t_{max}

The fact that $D^t_{max} \geq D^{star}_{max}$ (Theorem 3) and $D^t_{min} \leq D^{star}_{min}$ [13, 20] gives

$$D^t_{max} - D^t_{min} \geq D^{star}_{max} - D^{star}_{min}$$
$$\Delta^t_{max} \geq \Delta^{star}_{max}$$

and also

$$\frac{D^t_{max}}{D^t_{min}} \geq \frac{D^{star}_{max}}{D^{star}_{min}}$$
$$\Gamma^t_{max} \geq \Gamma^{star}_{max}.$$

4.2. The maxima of Δ^t_{max} and Γ^t_{max}

The fact that $D^t_{max} \leq D^{b-bistar}_{max}$ (Theorem 2) and $D^t_{min} \geq D^{linear}_{min}$ [20] imply that

$$\Delta^t_{max} \leq D^{b-bistar}_{max} - D^{linear}_{min} \tag{34}$$

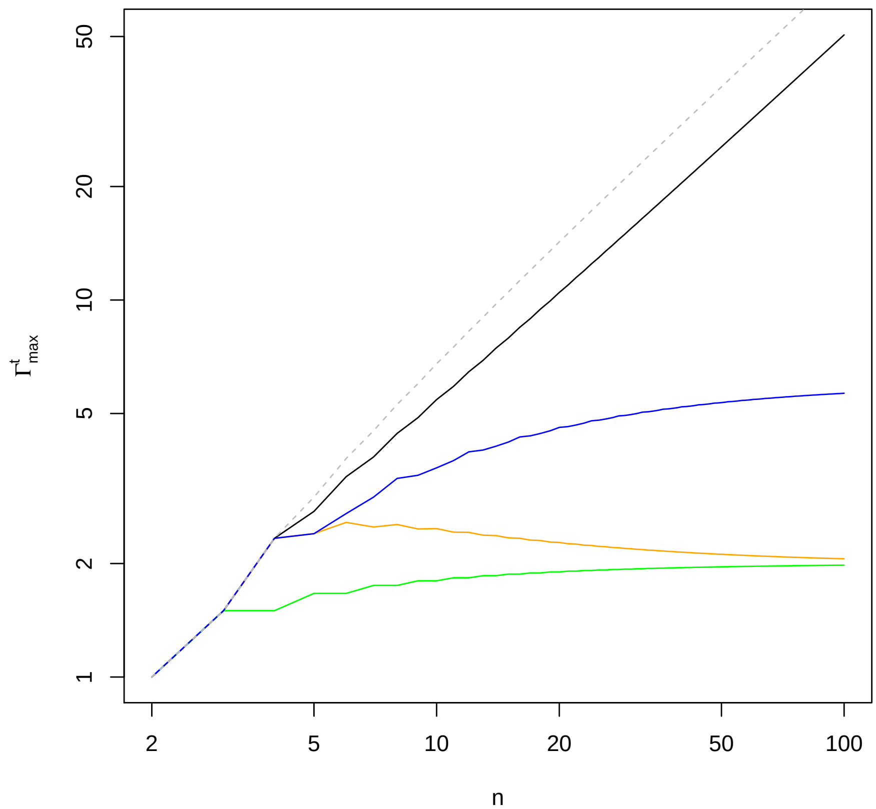


Figure 9. The scaling of Γ^t_{max} as a function of n, the number of vertices of the tree t for different trees: linear trees (black), quasistar trees (orange), star trees (green) and balanced bistar trees (blue). For reference, the upper bound $D^{b-bistar}_{max}/D^{linear}_{min}$ (dashed gray line) is also shown.

and also

$$\Gamma^t_{max} \leq \frac{D^{b-bistar}_{max}}{D^{linear}_{min}}. \tag{35}$$

However, these are unlikely to be tight upper bounds of Δ^t_{max} and Γ^t_{max} because the two kinds of trees involved in equation 34 and equation 35, star trees and balanced bistar trees, are not the same, contrary to what happened for the minima of Δ^t_{max} and Γ^t_{max}, given exactly by a star tree in both cases.

We perform a computational analysis of the maxima of Δ^t_{max} and Γ^t_{max} (the methods are explained in Appendix B). One the one hand, such analysis indicates that (Table 2)

$$\Delta^t_{max} \leq \Delta^{b-bistar}_{max} \tag{36}$$

for $n \leq 8$, consistently with figure 8, but

$$\Delta^t_{max} \leq \Delta^{t^*}_{max} \tag{37}$$

Table 2. Maximum Δ^t_{max} as a function of n and statistical properties of the trees that reach it: the kind of tree, $K_2 = n\langle k^2 \rangle$, the sum of squared degrees, k_1, the maximum degree, n_1, the number of leaves, L^t, the diameter in edges, $\langle l \rangle^t$, the average path length, and finally, D^t_{min} and D^t_{max}, the minimum and the maximum of D^t over all $n!$ linear arrangements. As for the kind of tree, *quasi* stands for quasi star tree, *b-bistar* for balanced bistar and *cat* for caterpillar.

n	$\Delta^{b-bistar}_{max}$	Maximum Δ^t_{max}	Kind of tree	K_2	k_1	n_1	L^t	$\langle l \rangle^t$	D^t_{min}	D^t_{max}
3	1	1	linear star b-bistar	6	2	2	2	1.33	2	3
4	4	4	linear quasi b-bistar	10	2	2	3	1.67	3	7
5	7	7	linear	14	2	2	4	2	4	11
			quasi b-bistar	16	3	3	3	1.8	5	12
6	12	12	linear	18	2	2	5	2.33	5	17
			cat	20	3	3	4	2.07	6	18
			b-bistar	22	3	4	3	1.93	7	19
7	18	18	cat	24	3	3	5	2.48	7	25
			cat	24	3	3	5	2.38	7	25
			cat	26	3	4	4	2.19	8	26
			cat	28	4	4	4	2.1	8	26
			b-bistar	30	4	5	3	2	9	27
8	26	26	cat	30	3	4	5	2.43	9	35
			cat	30	3	4	5	2.64	9	35
			cat	30	3	4	5	2.5	9	35
			cat	34	4	5	4	2.21	10	36
			b-bistar	38	4	6	3	2.07	11	37
9	34	35	cat	38	4	5	5	2.5	11	46
			cat	38	4	5	5	2.56	11	46
			cat	38	4	5	5	2.44	11	46
			cat	38	4	5	5	2.72	11	46
			cat	42	4	6	4	2.28	12	47
10	44	46	cat	46	4	6	5	2.47	13	59
			cat	46	4	6	5	2.82	13	59
			cat	46	4	6	5	2.56	13	59
11	55	57	cat	50	4	6	6	2.73	14	71
			cat	50	4	6	6	2.87	14	71
			cat	50	4	6	6	3.02	14	71
			cat	52	4	7	5	2.58	15	72
			cat	52	4	7	5	2.73	15	72
			cat	52	4	7	5	2.84	15	72
			cat	52	4	7	5	2.62	15	72
			cat	56	5	7	5	2.55	16	73
			cat	56	5	7	5	2.58	16	73
			cat	56	5	7	5	2.87	16	73
			cat	56	5	7	5	2.47	16	73

Table 3. Maximum Γ^t_{max} as a function of n and statistical properties of the trees that reach it. The format is based on that of Table 2. $\langle l \rangle^t = (n+1)/3$ as expected for a linear tree [4].

n	Γ^{linear}_{max}	Maximum Γ^t_{max}	Kind of tree	K_2	k_1	n_1	L^t	$\langle l \rangle^t$	D^t_{min}	D^t_{max}
3	1.5	1.5	linear star b-bistar	6	2	2	2	1.33	2	3
4	2.33	2.33	linear quasi b-bistar	10	2	2	3	1.67	3	7
5	2.75	2.75	linear	14	2	2	4	2	4	11
6	3.4	3.4	linear	18	2	2	5	2.33	5	17
7	3.83	3.83	linear	22	2	2	6	2.67	6	23
8	4.43	4.43	linear	26	2	2	7	3	7	31
9	4.88	4.88	linear	30	2	2	8	3.33	8	39
10	5.44	5.44	linear	34	2	2	9	3.67	9	49
11	5.9	5.9	linear	38	2	2	10	4	10	59

for $9 \leq n \leq 11$, where t^* is some caterpillar tree that is neither a bistar nor a linear tree. In addition, the bistar tree is not the only tree maximizing Δ^t_{max} for $4 \leq n \leq 8$ (Table 2). Notice that, for $n = 3$, a linear tree, a star tree and a balanced bistar tree are actually the same tree (when $n = 4$, the linear tree and the balanced bistar tree are the same tree). On the other hand, it indicates that (Table 3)

$$\Gamma^t_{max} \leq \Gamma^{linear}_{max} \tag{38}$$

for $n \leq 11$, consistently with figure 9. Interestingly, the linear tree is the only tree maximizing Γ^t_{max} up to $n = 11$ (Table 3).

4.3. The relationship with Δ_{rla} and Γ_{rla}

We define the expected value of Δ^t and Γ^t in a random linear arrangement (rla) of a given tree t as Δ^t_{rla} and Γ^t_{rla} respectively. Recall that $D_{rla} = \mathbb{E}_{rla}[D]$. Given a tree t, D^t_{min} and D_{rla} are constant and then

$$\begin{aligned}
\Delta^t_{rla} &= \mathbb{E}_{rla}[\Delta^t] \\
&= \mathbb{E}_{rla}[D^t - D^t_{min}] \\
&= D_{rla} - D^t_{min} \\
\Gamma^t_{rla} &= \mathbb{E}_{rla}[\Gamma^t] \\
&= \mathbb{E}_{rla}\left[\frac{D^t}{D^t_{min}}\right] \\
&= \frac{D_{rla}}{D^t_{min}}.
\end{aligned}$$

The fact that

$$D^t_{min} \leq D_{rla} \leq D^t_{max}$$

Table 4. $\langle k^2 \rangle$, the second moment of degree about zero and $\mathbb{V}^t_{rla}$, the variance of D^t in uniformly random linear arrangements of the tree t, in specific trees. $a = \frac{1}{180}(n+1)$. $\langle k^2 \rangle$ for linear and star trees is borrowed from [14]. $\langle k^2 \rangle$ for quasistar trees is borrowed from [23] and that of balanced bistar trees is derived from equation 10 with $k_1 = \lceil n/2 \rceil$. $\mathbb{V}^{linear}_{rla}$ and $\mathbb{V}^{star}_{rla}$ are borrowed from [17]. $\mathbb{V}^{b-bistar}_{rla}$ and $\mathbb{V}^{quasi}_{rla}$ are derived from 39 and the corresponding value of $\langle k^2 \rangle$ in this table.

t	$\langle k^2 \rangle$	$\mathbb{V}^t_{rla}$
linear	$4 - \frac{6}{n}$	$\frac{1}{90}(n-2)(n+1)(4n-7)$
balanced bistar	$\frac{2}{n}(\lceil n/2 \rceil(\lceil n/2 \rceil - n) - 1)$ $+n+1$	$a\left[2(n-4)\lceil n/2 \rceil(\lceil n/2 \rceil - n) + n(n^2-n-14) + 12\right]$
quasistar	$n - 3 + \frac{6}{n}$	$a[n((n-3)n+10) - 20]$
star	$n-1$	$a(n-1)(n+2)(n-2)$

gives

$$0 = \Delta^t_{min} \leq \Delta^t_{rla} \leq \Delta^t_{max}$$
$$1 = \Gamma^t_{min} \leq \Gamma^t_{rla} \leq \Gamma^t_{max}.$$

5. The minimum and the maximum z-score

For a specific tree t of n vertices, the minimum and the maximum values of D^t_z over all possible linear arrangements are

$$D^t_{z,min} = \frac{D^t_{min} - D_{rla}}{(\mathbb{V}^t_{rla})^{1/2}}$$
$$D^t_{z,max} = \frac{D^t_{max} - D_{rla}}{(\mathbb{V}^t_{rla})^{1/2}}.$$

Table 1 and [17]

$$\mathbb{V}^t_{rla} = \frac{n+1}{45}\left[(n-1)^2 + \left(\frac{n}{4} - 1\right) n \langle k^2 \rangle\right] \tag{39}$$

allow one to obtain formulae of $D^t_{z,min}$ and $D^t_{z,max}$ for specific trees. Table 4 summarizes $\mathbb{V}^t_{rla}$ in these trees. Let us consider $D^{linear}_{z,min}$ as an example. The numerator of $D^{linear}_{z,min}$ is (Table 1 and equation 3)

$$D^{linear}_{min} - D_{rla} = -\frac{1}{3}(n-1)(n-2) \leq D - D_{rla}.$$

whereas the denominator is $\mathbb{V}^{linear}_{rla}$ (Table 4). Then

$$D^{linear}_{z,min} = -(n-1)\left[\frac{10(n-2)}{(n+1)(4n-7)}\right]^{-1/2}.$$

Figures 10 and 11 show the evolution of D^t_{min} and $D^t_{z,max}$ as n increases for specific trees.

In star trees, quasistar trees and balanced bistar trees, both $D^t_{min} - D_{rla}$, $D^t_{max} - D_{rla}$ and $\mathbb{V}^t_{rla}$ are quadratic functions of n (Tables 1 and 4). In balanced bistar trees, the leading coefficient of $D^t_{min} - D_{rla}$ is $1/8 - 1/3 = -5/24$ and that of $\mathbb{V}^t_{rla}$ is $1/(6\sqrt{10})$, giving

$$\lim_{n\to\infty} D^{b-bistar}_{z,min} = -\frac{5}{4}\sqrt{10}.$$

In stars and quasistars, the leading coefficients are $1/4 - 1/3 = -1/12$ and $1/(6\sqrt{5})$. Hence

$$\lim_{n\to\infty} D^{star}_{z,min} = \lim_{n\to\infty} D^{quasi}_{z,min} = -\frac{\sqrt{5}}{2}.$$

In balanced bistar trees, the leading coefficient of $D^t_{max} - D_{rla}$ is $3/4 - 1/3 = 5/12$ while that of $\mathbb{V}^t_{rla}$ is $1/(6\sqrt{10})$, giving

$$\lim_{n\to\infty} D^{b-bistar}_{z,max} = \frac{5}{2}\sqrt{10}.$$

In stars and quasistars, the leading coefficients are $1/2 - 1/3 = 1/6$ and $1/(6\sqrt{5})$. Hence

$$\lim_{n\to\infty} D^{star}_{z,max} = \lim_{n\to\infty} D^{quasi}_{z,max} = \sqrt{5}.$$

These limiting values are consistent with figures 10 and 11.

5.1. The minima and the maxima of $D^t_{z,min}$

Equation 4 in combination with [17]

$$\mathbb{V}^{linear}_{rla} \leq \mathbb{V}^t_{rla} \leq \mathbb{V}^{star}_{rla} \tag{40}$$

yield that

$$D^{linear}_{z,min} \leq D^t_{z,min} \leq D^{star}_{z,min}. \tag{41}$$

5.2. The minima and the maxima of $D^t_{z,max}$

$D^{star}_{max} \leq D^t_{max} \leq D^{b-bistar}_{max}$ (Theorems 2 and 3) in combination with equation 40 yield

$$D^{star}_{z,min} \leq D^t_{z,max} \leq \frac{D^{b-bistar}_{max} - D_{rla}}{\mathbb{V}^{linear}_{rla}}. \tag{42}$$

Again, the latter upper bound is unlikely to be a tight upper bound of $D^t_{z,max}$ because the two kinds of trees involved (a balanced bistar and a linear tree) are not the same. Interestingly, figure 11 shows that $D^{bistar}_{z,max} \leq D^{linear}_{z,max}$ only for $n < 82$.

The computational analysis (Appendix B) in Table 5 indicates that

$$D^t_{z,max} \leq D^{b-bistar}_{z,max} \tag{43}$$

for $n \leq 10$, consistently with figure 11. The balanced bistar tree is the only tree maximizing $D^t_{z,max}$ up to $n = 10$ (Table 5). Contrary to expectations, the trend is broken for $n = 10$ because

$$D^t_{z,max} \leq D^{t^*}_{z,max}$$

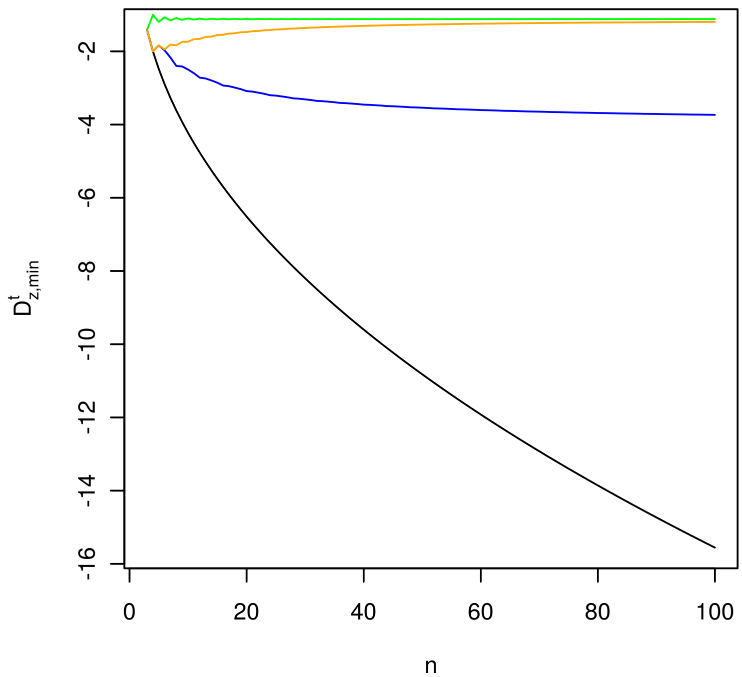


Figure 10. The scaling of $D^t_{z,min}$ as a function of n, the number of vertices of the tree t for different trees: linear trees (black), quasistar trees (orange), star trees (green) and balanced bistar trees (blue).

Table 5. Maximum $D^t_{z,max}$ as a function of n and statistical properties of the trees that reach it. The format is based on that of Table 2.

n	$D^{b-bistar}_{z,max}$	Maximum $D^t_{z,max}$	Kind of tree	K_2	k_1	n_1	L^t	$\langle l \rangle^t$	D^t_{min}	D^t_{max}
3	0.71	0.71	linear star b-bistar	6	2	2	2	1.33	2	3
4	2	2	linear quasi b-bistar	10	2	2	3	1.67	3	7
5	2.45	2.45	quasi b-bistar	16	3	3	3	1.8	5	12
6	3.1	3.1	b-bistar	22	3	4	3	1.93	7	19
7	3.41	3.41	b-bistar	30	4	5	3	2	9	27
8	3.84	3.84	b-bistar	38	4	6	3	2.07	11	37
9	4.06	4.06	b-bistar	48	5	7	3	2.11	14	48
10	4.37	4.37	b-bistar	58	5	8	3	2.16	17	61
11	4.54	4.56	cat	62	5	8	4	2.33	18	74

with t^* being some caterpillar tree that is neither a bistar nor a linear tree. That tree has only three internal vertices and consists of a star tree with a hub of degree $k_1 = 6$ that is attached to two stars whose hubs have the same degree ($k_2 = k_3 = 3$).

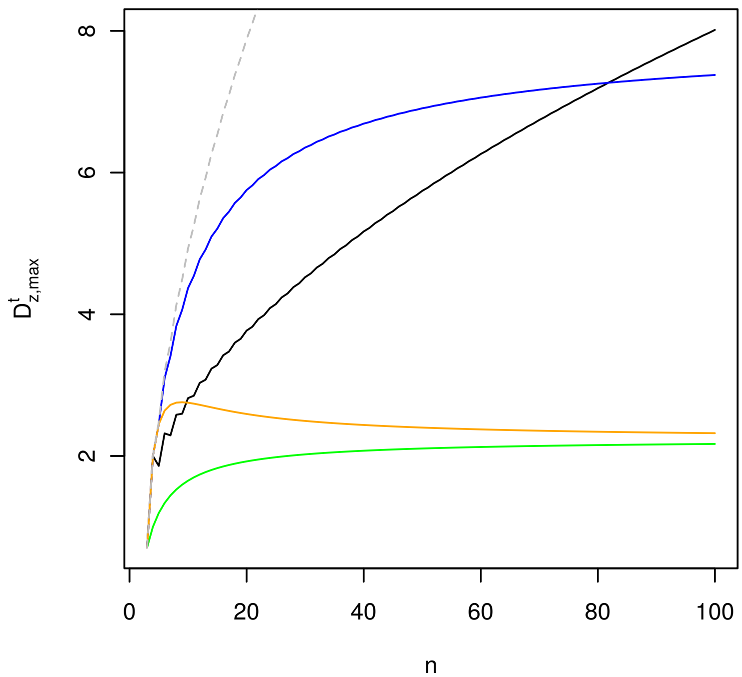


Figure 11. The scaling of $D^t_{z,max}$ as a function of n, the number of vertices of the tree t for different trees: linear trees (black), quasistar trees (orange), star trees (green) and balanced bistar trees (blue). For reference, the upper bound $(D^{b-bistar}_{max} - D_{rla})/\mathbb{V}^{linear}_{rla}$ (dashed gray line) is also shown.

5.3. The relationship with $D_{z,rla}$

We define $D^t_{z,rla}$ as the expected the value of D^t_z in a random linear arrangement (rla) of a given tree t. As D^t_{min}, D_{rla} and $\mathbb{V}^t_{rla}$ are constant given a tree t, one has

$$\begin{aligned} D^t_{z,rla} &= \mathbb{E}_{rla}\left[\frac{D^t - D_{rla}}{(\mathbb{V}^t_{rla})^{1/2}}\right] \\ &= \frac{\mathbb{E}_{rla}[D^t] - D_{rla}}{(\mathbb{V}^t_{rla})^{1/2}} \\ &= 0. \end{aligned}$$

The fact that

$$D^t_{min} \leq D_{rla} \leq D^t_{max}$$

gives

$$D^t_{z,min} \leq D^t_{z,rla} \leq D^t_{z,max}.$$

6. Discussion

The main results of the preceding sections have been validated using a computational procedure described in the Appendix B.

We have investigated the limits of the variation of D^t, the sum of edge lengths of trees of a given size n (Table 1). As for D^t_{min}, we have contributed with new formulae for the class of caterpillar trees that depend only on n and the vertex degrees, complementing the pioneering research in [29]. These formulae have allowed us to obtain formulae for the subclass of bistar trees that depend only on n and k_1, the maximum degree, which in turn have allowed us to obtain new formulae that depend only on n for specific trees: quasistar trees and balanced bistar trees. [29] obtained a lower bound for D^t_{min} (Table 1) that gives actually the exact value of D^t_{min} when t is a caterpillar. We have contributed with a much shorter proof of the argument and showing that the lower bound is actually a significant improvement with respect to previous attempts to provide a lower bound of D^t_{min} based on vertex degrees [38, 14]. Therefore, although D^t_{min} can be calculated in polynomial time employing existing algorithms [10, 11, 12], D^t_{min} can be calculated in constant time for caterpillar given trees of size n, $\langle k^2 \rangle$ and q (Table 1).

As for D^t_{max}, we have not found a simple enough formula for the class of caterpillar trees but we have obtained one for the subclass of bistar trees as function of n and k_1. Thanks to this work we have obtained new formulae that depend only on n for specific trees: quasistar trees and balanced bistar trees (Table 1). The new formula of D^t_{max} for linear trees has been obtained employing an independent analysis. A unified derivation of D^t_{max} for linear trees and bistar trees, as well as a general but simple formula of D^t_{max} for caterpillar trees, should be the subject of future research. Finally, we delimited the range of variation of D^t_{max}, obtaining the following chain of inequalities

$$D_{rla} \leq D^{star}_{max} \leq D^t_{max} \leq D^{b-bistar}_{max}. \tag{44}$$

The importance of this chain is two-fold. First, it indicates that the problem of maximizing D^g and that of minimizing D^g are not symmetric, because the corresponding chain for the minimization problem does not involve balanced bistar trees (equation 4). Second, it links the problem of maximizing D^t without constraints (i.e. D^t_{max}) with the problem of maximizing D^t under the planarity constraint (i.e. $D^t_{max,P}$), since $D^t_{max,P} \leq D^{linear}_{max,P} = D^{star}_{max}$ [14]. The finding indicates that any tree has a linear arrangement reaching the maximum possible D^t for any tree under the planarity constraint, namely $D^t_{max} \geq D^{linear}_{max,P} = D^{star}_{max}$ ([14] did not address the question of whether $D^t_{max,P} = D^{linear}_{max,P} = D^{star}_{max}$ for any other tree t). Real syntactic dependency trees are almost planar in the sense that edge crossings are scarce [44] and the origin of such a characteristic is being debated [45].

In this article, we have established some mathematical foundations for the analysis and development of optimality scores based on D^t and explored some implications for the limits of the variation of two scores: Γ^t and Δ^t. We have obtained the following chains of inequalities:

$$0 = \Delta^t_{min} \leq \Delta^t_{rla} \leq \Delta^{star}_{max} \leq \Delta^t_{max} \tag{45}$$

$$1 = \Gamma^t_{min} \leq \Gamma^t_{rla} \leq \Gamma^{star}_{max} \leq \Gamma^t_{max}. \tag{46}$$

We conjecture that $\Gamma^t_{max} \leq \Gamma^{linear}_{max}$ and that the linear tree is the only maximum of Γ^t_{max} (Table 3). A linear tree is the tree that minimizes the denominator of Γ^t_{max}. The numerator is maximized by a balanced bistar tree but it is easy to show (just using the formulae in Table 1) that $\Gamma^t_{max} \leq \Gamma^{linear}_{max}$ for any tree t that is a bistar. Similarly, we have obtained the following chains of inequalities for the z-score:

$$D^{linear}_{z,min} \leq D^t_{z,min} \leq D^{star}_{z,min} \leq D_{z,rla} = 0 \tag{47}$$

$$0 = D_{z,rla} \leq D^{star}_{z,max} \leq D^t_{z,max}. \tag{48}$$

The problem of the trees that maximize Γ^t_{max}, Δ^t_{max} and $D^t_{z,max}$ should receive further investigation in two directions: characterizing the trees that maximize these scores (proving or refuting the conjectures above) or, at least, expanding the range of n for which the true optima are known. We hope that our findings stimulate further research on optimality scores in linear arrangements.

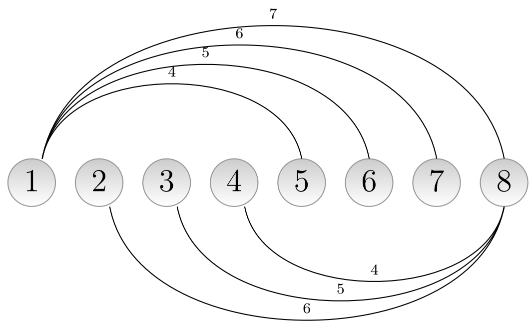
Acknowledgments

We are very grateful to L. Alemany-Puig for his careful revision of the manuscript. JLE is funded by the grant TASSAT3 (TIN2016-76573-C2-1-P) from MINECO (Ministerio de Economia, Industria y Competitividad). CGR is funded by the European Research Council (ERC), under the European Union's Horizon 2020 research and innovation programme (FASTPARSE, grant agreement No 714150), the ANSWER-ASAP project (TIN2017-85160-C2-1-R) from MINECO and Xunta de Galicia (ED431B 2017/01, ED431G2019/01, and an Oportunius program grant to complement ERC grants). RFC is supported by the grant TIN2017-89244-R from MINECO (Ministerio de Economia, Industria y Competitividad) and the recognition 2017SGR-856 (MACDA) from AGAUR (Generalitat de Catalunya).

Appendix A. The upper bound of D^t_{max}

Suppose that vertices are labelled with positions in the linear arrangement. An edge between vertices i and j is indicated by the unordered pair $\{i, j\}$. The problem of obtaining a tree that maximizes D^t_{max} for any tree t of n vertices is equivalent to the problem of finding the maximum spanning tree of a complete graph where the weight of the edge $\{i, j\}$ is $|i - j|$, as each possible spanning tree bijectively corresponds to a linear arrangement of some tree of n vertices, and the sum of weights corresponds to its value of D. We will show that a balanced bistar tree is the outcome of an algorithm that is based on Prim's algorithm to find the minimum spanning tree of a graph [46]. Prim's original algorithm solves a minimization problem. The maximization problem

(a)



(b)

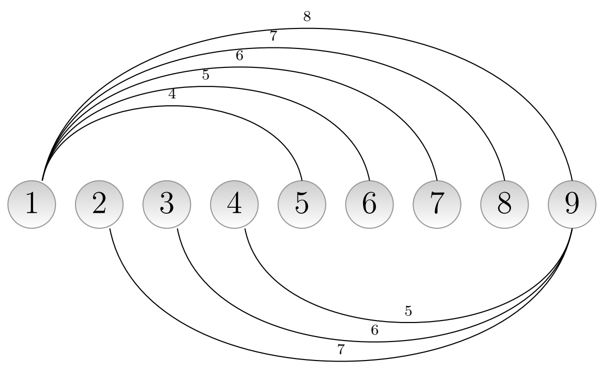


Figure A1. Linear arrangements of balanced bistar trees of n vertices that maximize D^t. Vertex labels indicate the position of each vertex. Edge labels indicate edge distances. (a) $n = 8$ and $D^t = 37$. (b) $n = 9$ and $D^t = 48$.

can be solved using the customary minimization version with edge weights defined as $n - |i - j|$. We use a variant of Prim's algorithm to solve the maximization problem that eases the proof:

(i) Initialize the tree t with vertex 1.

(ii) Find the edge linking one vertex in t and another vertex outside t such that the weight is maximized. Add the edge (and the new vertex) to t.

(iii) Repeat step ii until t has n vertices.

In the context of our application, i.e. the maximization of D for any possible tree t of n vertices, this variant of Prim's algorithm becomes

(i) Initialize the tree t with vertex 1.

(ii) Set x to 2 and y to n.

(iii) Compare the length of the edges $\{1, y\}$ and $\{x, n\}$. If the longest edge is $\{1, y\}$, add the edge (and vertex y) to t and decrement y. Otherwise, add $\{x, n\}$ (and add x) to t and increment x.

(iv) Repeat step iii until t has n vertices.

Iteration	Edge	Length	$[x, y]$
0	-	-	$[2, n]$
1	$\{1, n\}$	$n - 1$	$[2, n - 1]$
2	$\{1, n - 1\}$	$n - 2$	$[2, n - 2]$
3	$\{2, n\}$	$n - 2$	$[3, n - 2]$
4	$\{1, n - 2\}$	$n - 3$	$[3, n - 3]$
5	$\{3, n\}$	$n - 3$	$[4, n - 3]$
...	...	...	...

Table A1. The edge added at every iteration, its length and $[x, y]$, the interval of vertex labels that do not belong to t after adding the edge.

Notice that the vertices that do not belong to t are in the interval $[x, y]$. As for Step iii, notice that the longest edge liking one vertex in t, namely one vertex in $[1, n] \setminus [x, y]$, and another vertex outside t, namely one vertex in $[x, y]$, can only be $\{1, y\}$ or $\{x, n\}$. It is easy to see that the execution of this algorithm produces edges that correspond to a balanced bistar tree (Table A1) that is arranged linearly as in Fig. A1.

Appendix B. Validation

The main results of the article, namely Table 1 and the chains of inequalities in equations 44, 45, 46, 47 and 48 have been validated using a brute force procedure up to $n = 11$ inspired by that of [20]. For a given n, the procedure calculates D^t_{min} and D^t_{max} for every distinct unlabelled tree and consists in generating all the n^{n-2} labelled trees using Prüfer codes as in [20] while updating a two-level table containing the current value of D^t_{min} and D^t_{min} and a signature of the tree to speed up the tree isomorphism test [47]. The signature of a tree is defined as a vector containing the canonical names [47] of the trees rooted at each of the Jordan centers [48] of the original free tree. A tree has 1 or 2 Jordan centers [48]. For each labelled tree whose underlying unlabelled tree is t,

(i) D^t is calculated interpreting vertex labels as vertex positions in the linear arrangement.

(ii) The signature of t for the test of tree isomorphism is calculated.

(iii) t is searched in the collection of already visited unlabelled trees. The unlabelled trees are accessed using a two-level look-up table: first, by their value of $n \langle k^2 \rangle$ and second, by the degree spectrum. The frequency spectrum is a vector indicating the number of vertices of that have a certain degree k. Then, the corresponding unlabelled tree is found comparing all the stored trees with the same degree spectrum against the target tree using using their respective signatures.

(iv) If t is new, then both D^t_{min} and D^t_{max} are set to D^t temporarily.

(v) If t is not new, then D^t_{min} and D^t_{max} are updated based on D^t.

At the end of the exploration, one has the exact value of D^t_{min} and D^t_{max} for every tree

t. As a sanity check, we verify that the number of labelled trees in the look-up table is the one expected by OEI A00055, `https://oeis.org/A000055`. We also verify, for every tree t, that

(i) D^t_{min} coincides with the value obtained by the corrected version of Shiloach's algorithm [11] as a sanity check.

(ii) D^t_{min} and D^t_{max} match the predictions in Table 1 and satisfy the inequalities in 44,45 and 46.

Equations 36, 37, 38 and 43 have been inferred using the procedure above.

References

[1] Vito Latora and Massimo Marchiori. Efficient behavior of small-world networks. *Phys. Rev. Lett.*, 87(19):198701, Oct 2001.

[2] R. Ferrer-i-Cancho. Euclidean distance between syntactically linked words. *Physical Review E*, 70:056135, 2004.

[3] M. Barthélemy. *Morphogenesis of Spatial Networks*. Springer, Cham, Switzerland, 2018.

[4] R. Ferrer-i-Cancho and R. V. Solé. Optimization in complex networks. In R. Pastor-Satorras, J.M. Rubí, and A. Díaz-Guilera, editors, *Statistical Mechanics of complex networks*, volume 625 of *Lecture Notes in Physics*, pages 114–125. Springer, Berlin, 2003.

[5] Gorka Zamora-López and Romain Brasselet. Sizing complex networks. *Communications Physics*, 2(1):144, 2019.

[6] V. Latora and M. Marchiori. Economic small-world behavior in weighted networks. *The European Physical Journal B - Condensed Matter and Complex Systems*, 32(2):249–263, 2003.

[7] H. Liu, C. Xu, and J. Liang. Dependency distance: A new perspective on syntactic patterns in natural languages. *Physics of Life Reviews*, 21:171–193, 2017.

[8] J. Díaz, J. Petit, and M. Serna. A survey of graph layout problems. *ACM Computing Surveys*, 34:313–356, 2002.

[9] R. Hassin and S. Rubinstein. Approximation algorithms for maximum linear arrangement. *Information Processing Letters*, 80(4):171 − 177, 2001.

[10] Y. Shiloach. A minimum linear arrangement algorithm for undirected trees. *SIAM J. Comput.*, 8(1):15–32, 1979.

[11] J. L. Esteban and R. Ferrer-i-Cancho. A correction on Shiloach's algorithm for minimum linear arrangement of trees. *SIAM Journal of Computing*, 46(3):1146–1151, 2015.

[12] F. R. K. Chung. On optimal linear arrangements of trees. *Comp. & Maths. with Appls.*, 10(1):43–60, 1984.

[13] M. A. Iordanskii. Minimal numberings of the vertices of trees. *Dokl. Akad. Nauk SSSR*, 218(2):272–275, 1974.

[14] R. Ferrer-i-Cancho. Hubiness, length, crossings and their relationships in dependency trees. *Glottometrics*, 25:1–21, 2013.

[15] Harry J. Tily. *The role of processing complexity in word order variation and change*. PhD thesis, Stanford University, August 2010. Chapter 3: Dependency lengths.

[16] Kristina Gulordava and Paola Merlo. Diachronic trends in word order freedom and dependency length in dependency-annotated corpora of Latin and ancient Greek. In *Proceedings of the Third International Conference on Dependency Linguistics (Depling 2015)*, pages 121–130, Uppsala, Sweden, 2015. Uppsala University.

[17] R. Ferrer-i-Cancho. The sum of edge lengths in random linear arrangements. *Journal of Statistical Mechanics*, page 053401, 2019.

[18] R. Ferrer-i-Cancho and C. Gómez-Rodríguez. Anti dependency length minimization in short sequences. a graph theoretic approach. *Journal of Quantitative Linguistics*, page in press. doi: 10.1080/09296174.2019.1645547, 2019.

[19] David Temperley and Daniel Gildea. Minimizing syntactic dependency lengths: Typological/Cognitive universal? *Annual Review of Linguistics*, 4(1):67–80, 2018.

[20] J. L. Esteban, R. Ferrer-i-Cancho, and C. Gómez-Rodríguez. The scaling of the minimum sum of edge lengths in uniformly random trees. *Journal of Statistical Mechanics*, page 063401, 2016.

[21] Immanuel T. San Diego and Frederick S. Gella. The b-chromatic number of bistar graph. *Applied Mathematical Sciences*, 8(116):5795–5800, 2014.

[22] Samir K. Vaidya and Sejal H. Karkar. Steiner domination number of splitting and degree splitting graphs. *International J. Math. Combin.*, 3:81–86, 2018.

[23] R. Ferrer-i-Cancho. Non-crossing dependencies: least effort, not grammar. In A. Mehler, A. Lücking, S. Banisch, P. Blanchard, and B. Job, editors, *Towards a theoretical framework for analyzing complex linguistic networks*, pages 203–234. Springer, Berlin, 2016.

[24] L. Alemany-Puig and R. Ferrer-i-Cancho. Edge crossings in random linear arrangements. *Journal of Statistical Mechanics*, page 023403, 2020. http://dx.doi.org/10.1088/1742-5468/ab6845.

[25] G. N. Frederickson and S. E. Hambrusch. Planar linear arrangements of outerplanar graphs. *IEEE Transactions on Circuits and Systems*, 35(3):323–333, 1988.

[26] Raghunath Raghavan and Sartaj Sahni. Optimal single row router. In *Proceedings of the 19th Design Automation Conference*, DAC 82, page 3845. IEEE Press, 1982.

[27] Carlos Gómez-Rodríguez and Joakim Nivre. Divisible transition systems and multiplanar dependency parsing. *Computational Linguistics*, 39(4):799–845, 2013.

[28] K.H. Rosen, D. R. Shier, and W. Goddard. *Hanbook of discrete and combinatorial mathematics*. CRC Press, Boca Raton, FL, 2017.

[29] S. Horton. *The optimal linear arrangement problem: algorithms and approximation*. PhD thesis, Georgia Institute of Technology, 1997.

[30] Y. Albert Park and R. Levy. Minimal-length linearizations for mildly context-sensitive dependency trees. In *Proceedings of the 10th Annual Meeting of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT) conference*, pages 335–343, Stroudsburg, PA, USA, 2009. Association for Computational Linguistics.

[31] R. Ferrer-i-Cancho and H. Liu. The risks of mixing dependency lengths from sequences of different length. *Glottotheory*, 5:143–155, 2014.

[32] R. Futrell, K. Mahowald, and E. Gibson. Large-scale evidence of dependency length minimization in 37 languages. *Proceedings of the National Academy of Sciences*, 112(33):10336–10341, 2015.

[33] Chern-Ching Chao and Wen-Qi Liang. Arranging n distinct numbers on a line or a circle to reach extreme total variations. *European Journal of Combinatorics*, 13(5):325–334, 1992.

[34] Feodor F. Dragan and Chenyu Yan. Distance approximating trees: Complexity and algorithms. In Tiziana Calamoneri, Irene Finocchi, and Giuseppe F. Italiano, editors, *Algorithms and Complexity*, pages 260–271, Berlin, Heidelberg, 2006. Springer Berlin Heidelberg.

[35] B. Bollobás. *Modern graph theory*. Springer-Verlag, 1998.

[36] M. Boguñá, R. Pastor-Satorras, and A. Vespignani. Epidemic spreading in complex networks with degree correlations. In R. Pastor-Satorras, J.M. Rubí, and A. Díaz-Guilera, editors, *Statistical Mechanics of complex networks*, volume 625 of *Lecture Notes in Physics*, pages 127–147. Springer, Berlin, 2003.

[37] R. Ferrer-i-Cancho. The placement of the head that minimizes online memory. A complex systems approach. *Language Dynamics and Change*, 5(1):114–137, 2015.

[38] J. Petit. Experiments on the minimum linear arrangement problem. *Journal of Experimental Algorithmics*, 8, 2003.

[39] M. A. Iordanskii. Minimal numberings of the vertices of trees — Approximate approach. In Lothar Budach, Rais Gatič Bukharajev, and Oleg Borisovič Lupanov, editors, *Fundamentals of*

Computation Theory, pages 214–217, Berlin, Heidelberg, 1987. Springer Berlin Heidelberg.

[40] R. A. Hochberg and M. F. Stallmann. Optimal one-page tree embeddings in linear time. *Information Processing Letters*, 87:59–66, 2003.

[41] Daniel Gildea and David Temperley. Optimizing grammars for minimum dependency length. In *Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics*, pages 184–191, Prague, Czech Republic, June 2007. Association for Computational Linguistics.

[42] F. Shahrokhi, O. Sýkora, L.A. Székely, and I. Vrto. On bipartite drawings and the linear arrangement problem. *SIAM Journal on Computing*, 30(6):1773–1789, 2001.

[43] M. Chimani, S. Felsner, S. Kobourov, T. Ueckerdt, P. Valtr, and A. Wolff. On the maximum crossing number. *Journal of Graph Algorithms and Applications*, 22(1):67–87, 2018.

[44] R. Ferrer-i-Cancho, C. Gómez-Rodríguez, and J. L. Esteban. Are crossing dependencies really scarce? *Physica A: Statistical Mechanics and its Applications*, 493:311–329, 2018.

[45] C. Gómez-Rodríguez, M.H. Christiansen, and Ramon Ferrer-i-Cancho. Memory limitations are hidden in grammar. *https://arxiv.org/abs/1908.06629*, page under review, 2020.

[46] Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, and Clifford Stein. *Introduction to Algorithms*. The MIT Press, Cambridge, MA, USA, 2nd edition, 2001.

[47] Douglas M. Campbell and David Radford. Tree isomorphism algorithms: Speed vs. clarity. *Mathematics Magazine*, 64(4):252–261, 1991.

[48] S. Mitchell Hedetniemi, E. J. Cockayne, and S. T. Hedetniemi. Linear algorithms for finding the Jordan center and path center of a tree. *Transportation Science*, 15(2):98–114, 1981.